# Chapter 1
# Quantum Information Theory for Quantum Communication

**Masato Koashi**

## 1.1 Basic Rules of Quantum Mechanics

We begin by listing a basic set of rules from which all statements in this section should be derived. The choice of this set is by no means unique, and the selection of the properties of quantum mechanics that are used as basic rules, leaving the rest as derived rules, is actually a matter of preference. Our choice here comprises five rules describing states, transformations, measurements, compositions, and causality.

The first of these rules covers the description of the states of a physical system. We call a state *pure* when it is impossible to regard that state as a probabilistic mixture of two or more different[1] states.

Rule 1 A physical system is associated with a Hilbert space $\mathscr{H}$. Every pure state of this system is represented by a normalized vector $|\phi\rangle \in \mathscr{H}$. For any normalized vector $|\psi\rangle \in \mathscr{H}$, it is possible to prepare the system in the state represented by $|\psi\rangle$.

To avoid complications, we assume in this section that the dimension $d = \dim \mathscr{H}$ of the Hilbert space is finite.[2] A physical system with a Hilbert space of dimension $d$ is often called a *d-level* system. Rule 1 dictates that any appropriate

[1] The operational meaning of two states being different is that a measurement exists on the physical system that can show the difference statistically.

[2] Rule 1 also implies that we exclude any cases where a physical law such as the superselection rule imposes an additional restriction on the preparable states.

M. Koashi (✉)
Photon Science Center, Graduate School of Engineering, The University of Tokyo, Bunkyo-ku,
Tokyo 113-8656, Japan
e-mail: koashi@qi.t.u-tokyo.ac.jp

Y. Yamamoto, K. Semba (eds.), *Principles and Methods of Quantum Information Technologies*, Lecture Notes in Physics 911, DOI 10.1007/978-4-431-55756-2_1

instruction for preparation of the physical system leads to either a pure state represented by a single vector $|\phi\rangle$ or a *mixed* state represented by an *ensemble* $\{(p_j, |\phi_j\rangle)\}$, which designates the situation where the system is prepared in state $|\phi_j\rangle$ with probability $p_j$. In either case, the representation is not unique: $|\phi\rangle$ and $e^{i\varphi}|\phi\rangle$ represent the same pure physical state. The different descriptions $\{(p_j, |\phi_j\rangle)\}$ and $\{(q_i, |\psi_i\rangle)\}$ may both refer to the same mixed state. We will introduce an alternative representation of the states, which is unique, in Sect. 1.2.

The next two rules cover the input-output relations of feasible operations on a physical system prepared in state $|\phi_{\text{in}}\rangle$. A state transformation refers to the case where the output is the quantum state $|\phi_{\text{out}}\rangle$ of the system after the operation. Rule 2 dictates the feasibility of *unitary transformations*, which are, in a sense, a basic set of transformations.

Rule 2 For any unitary operator $\hat{U}$ on $\mathscr{H}$, it is possible to implement a state transformation where every input state $|\phi_{\text{in}}\rangle \in \mathscr{H}$ evolves into state $|\phi_{\text{out}}\rangle = \hat{U}|\phi_{\text{in}}\rangle$.

When the output is a classical variable, we are then referring to a measurement. Rule 3 covers a basic set of measurements called *(complete) orthogonal measurements*.

Rule 3 For any orthonormal basis $\{|u_j\rangle\}_{j=1,\ldots,d}$ of $\mathscr{H}$, it is possible to implement a measurement that produces the outcome $j = 1, \ldots d$ with probability $p_j = |\langle u_j|\phi_{\text{in}}\rangle|^2$ when the system is in state $|\phi_{\text{in}}\rangle \in \mathscr{H}$ before the measurement is performed.

In this rule, we are not interested in the state of the measured system after the measurement is performed. The two rules above only refer to the feasibility of the limited sets of transformations and measurements. In general, a much wider variety of operations should be available on a physical system, and we will see the whole landscape of these operations in Sect. 1.4.

The next rule is a very special rule that allows us to weave the threads of Rules 1, 2 and 3 into a texture of quantum information with dazzling patterns and colors. This rule tells us how to apply the three rules above when dealing with multiple physical systems. Consider two physical systems, $A$ and $B$, which are *independently* accessible. For example, the two systems are well separated in space, meaning that one can freely operate on system $A$ without affecting system $B$ at all. We may call this type of operation *local*. In this case, we can treat the whole of systems $A$ and $B$ together as a single physical system (a *composite system AB*), or can focus on one of the two systems (a *subsystem*) with no interest in the other. Rule 4 provides the connection between these different viewpoints.

Rule 4 Suppose that the subsystems $A$ and $B$ are associated with the Hilbert spaces $\mathscr{H}_A$ and $\mathscr{H}_B$, respectively. The composite system $AB$ is then associated with a tensor-product space $\mathscr{H}_{AB} = \mathscr{H}_A \otimes \mathscr{H}_B$. Local operations (e.g., state preparations, state transformations, measurements) are represented by the appropriate tensor products.

Specifically, preparation of system $A$ in state $|\phi\rangle_A \in \mathscr{H}_A$ and system $B$ in state $|\psi\rangle_B \in \mathscr{H}_B$ is equivalent to the preparation of a composite system $AB$ in state

$|\phi\rangle_A \otimes |\psi\rangle_B \in \mathscr{H}_{AB}$. The state that can be written in this form is called a *product state*, and is often abbreviated as $|\phi\rangle_A|\psi\rangle_B$ or even $|\phi\psi\rangle_{AB}$. The unitary transformations $\hat{U}_A$ on system $A$ and $\hat{V}_B$ on $B$ result in the unitary transformation $\hat{U}_A \otimes \hat{V}_B$ on the composite system $AB$. Performing an orthogonal measurement with basis $\{|u_i\rangle_A\}_{i=1,\dots,d}$ on system $A$ and another with basis $\{|v_j\rangle_B\}_{j=1,\dots,d'}$ on system $B$ can be regarded as the performance of a single orthogonal measurement, where the outcome is represented by two numbers $(i, j)$, carried out on the composite system $AB$ with the orthonormal basis $\{|u_i\rangle_A \otimes |v_j\rangle_B\}_{j=1,\dots,d'}^{i=1,\dots,d}$ of $\mathscr{H}_{AB}$.

According to Rule 1, we should be able to prepare a state represented by any vector $|\Psi\rangle_{AB} \in \mathscr{H}_{AB}$, possibly by the tailoring of suitable interaction between systems $A$ and $B$. These vectors include, for example, $(|u_1\rangle_A|v_1\rangle_B + |u_2\rangle_A|v_2\rangle_B)/\sqrt{2}$, which can never be written in the form $|\phi\rangle_A \otimes |\psi\rangle_B$. This type of state is called *entangled*. Similarly, a unitary operator $\hat{U}_{AB}$ acting on $\mathscr{H}_{AB}$ is not necessarily a product $\hat{U}_A \otimes \hat{V}_B$, and the corresponding *global* unitary transformation should be feasible. There are also global orthogonal measurements, for which the orthonormal basis is composed of entangled state vectors.

Since the state of a composite system is not necessarily written as a product form, the definition of 'the state of a subsystem' is something of a moot point. Here, we adopt a definition with a clear operational meaning, called the *marginal* state of a subsystem, which is simply the state that the subsystem would be in if we discard all the other constituent subsystems. With regard to the marginal states, we assume the following.

Rule 5 The marginal state of a subsystem is not changed by operating on other subsystems, as long as no information on the outcome of the operation is referred.

This rule is expected to hold because there would otherwise be a test on system $A$ alone that would give clues on what operations were performed on a remote system $B$ without any communication between them. The rule sets a limitation on the physically allowed state transformations and measurements, which complements the fact that Rules 2 and 3 merely dictate what we can at least do.

## 1.2 Density Operators

In classical mechanics, a mixed state is simply regarded as a way to formulate an observer's lack of knowledge of the true state of a system. In principle, it is always possible to assume that there is an omnipotent observer who knows the exact state (the pure state) of every system. In quantum mechanics, however, this simple picture does not hold. When a composite system is in a pure state $|\Psi\rangle_{AB}$, we cannot associate the state of the subsystem $A$ with a single vector $|\phi\rangle_A \in \mathscr{H}_A$ unless $|\Psi\rangle_{AB}$ is a product state. Therefore, it is not always possible to assume that every system is in a pure state at the same time. In this subsection, we determine how we can represent the state of a subsystem when it is a part of a composite system in a pure state $|\Psi\rangle_{AB}$. We will see that the intuitive representation using an ensemble $\{(p_j, |\phi_j\rangle)\}$ is redundant in the sense that different descriptions may refer to the same physical state. This motivates us to introduce a density operator to offer a better representation in this respect. By using a helpful property of bipartite pure states called Schmidt decomposition, we will show that there is a one-to-one correspondence between the density operators and the physical states.

### *1.2.1 Measurement on a Subsystem*

Suppose that the composite system $AB$ is initially prepared in a pure state $|\Psi\rangle_{AB}$, and an orthogonal measurement with a basis $\{|v_j\rangle_B\}_{j=1,\ldots,d'}$ is then conducted on subsystem $B$, producing an outcome $j$ with a probability $p_j$. Let us derive a rule to calculate $p_j$ and identify the state of the subsystem $A$ that is conditioned on the value of $j$.

Our strategy is to observe what happens if we perform a measurement with arbitrary basis $\{|u_i\rangle_A\}_{i=1,\ldots,d}$ on system $A$. Regardless of the temporal order of the measurements on $A$ and $B$, Rules 3 and 4 dictate that the joint probability of the two outcomes $(i,j)$ is given by $p_{i,j} = |({}_A\langle u_i| \otimes {}_B\langle v_j|)|\Psi\rangle_{AB}|^2$. Let us introduce the unnormalized vector $|\tilde{\phi}_j\rangle_A := {}_B\langle v_j||\Psi\rangle_{AB} \in \mathscr{H}_A$. We then have $p_{i,j} = |{}_A\langle u_i|\tilde{\phi}_j\rangle_A|^2$ and $p_j = \sum_{i=1}^d p_{i,j} = \sum_{i=1}^d |{}_A\langle u_i|\tilde{\phi}_j\rangle_A|^2 = |{}_A\langle\tilde{\phi}_j|\tilde{\phi}_j\rangle_A|^2$. Using a normalized vector $|\phi_j\rangle_A := |\tilde{\phi}_j\rangle_A/\sqrt{p_j}$, we obtain an expression for the conditional probability, $p_{i|j} := p_{i,j}/p_j = |{}_A\langle u_i|\phi_j\rangle_A|^2$. Because the choice of the basis $\{|u_i\rangle_A\}_{i=1,\ldots,d}$ was arbitrary, comparison of this relationship to Rule 3 shows that the state of the subsystem $A$ conditioned on the outcome $j$ must be a pure state, which is represented by the vector $|\phi_j\rangle_A$. Noting that the measurement on $A$ can be performed immediately after the preparation of $|\Psi\rangle_{AB}$, we arrive at the following theorem.

**Theorem 1.** *Suppose that a composite system $AB$ is initially prepared in a pure state $|\Psi\rangle_{AB}$, and that an orthogonal measurement with a basis $\{|v_j\rangle_B\}_{j=1,\ldots,d'}$ is performed on subsystem $B$. The outcome $j$ then occurs with probability $p_j$ and, conditioned on $j$, the subsystem $A$ behaves as if it was initially prepared in the pure state $|\phi_j\rangle_A$, where*

$$\sqrt{p_j}|\phi_j\rangle_A = {}_B\langle v_j||\Psi\rangle_{AB} \tag{1.1}$$

*holds.*

### *1.2.2 Marginal State of a Subsystem*

The argument in the previous subsection immediately provides a description of the marginal state of the subsystem $A$ when the composite system $AB$ is prepared in the pure state $|\Psi\rangle_{AB}$. If the value of the outcome $j$ of the measurement on subsystem $B$

is unavailable, then the state of system $A$ after the measurement can be described by the ensemble $\{(p_j, |\phi_j\rangle_A)\}_{j=1,\dots,d'}$, where the probabilities $\{p_j\}$ and the vectors $\{|\phi_j\rangle_A\}$ are calculated from Eq. (1.1). From Rule 5, we see that the marginal state of subsystem $A$ before the measurement was performed is also $\{(p_j, |\phi_j\rangle_A)\}_{j=1,\dots,d'}$.

On the one hand, this description is helpful because it is sufficient to allow calculation of the statistics of the outcomes of further operations on system $A$ alone. On the other hand, the argument above also shows that the description of a mixed state by the ensemble is by no means unique. If we change the basis $\{|v_j\rangle_B\}_{j=1,\dots,d'}$ of the measurement to another basis, then the description of the state $\{(p_j, |\phi_j\rangle_A)\}_{j=1,\dots,d'}$ also changes through Eq. (1.1). This new ensemble should also be a valid representation of the same state.

**Lemma 1.** *Two ensembles,* $\{(p_j, |\phi_j\rangle_A)\}_{j=1,\dots,d'}$ *and* $\{(p'_j, |\phi'_j\rangle_A)\}_{j=1,\dots,d'}$*, represent the same mixed state if a bipartite pure state* $|\Psi\rangle_{AB}$ *and orthonormal bases* $\{|v_j\rangle_B\}_{j=1,\dots,d'}$ *and* $\{|v'_j\rangle_B\}_{j=1,\dots,d'}$ *exist that satisfy*

$$\sqrt{p_j}|\phi_j\rangle_A = {}_B\langle v_j||\Psi\rangle_{AB} \text{ and } \sqrt{p'_j}|\phi'_j\rangle_A = {}_B\langle v'_j||\Psi\rangle_{AB}. \tag{1.2}$$

### 1.2.3 Density Operators

Consider a physical system that is associated with a Hilbert space $\mathscr{H}$, and let us call an operator $\hat{\rho} : \mathscr{H} \to \mathscr{H}$ a *density operator* when it is positive ($\hat{\rho} \geq 0$) and of unit trace ($\mathrm{Tr}\,\hat{\rho} = 1$). We associate a mixed state of a system represented by the ensemble $\{(q_i, |\psi_i\rangle)\}_{i=1,\dots,n}$ with a density operator given by

$$\hat{\rho} := \sum_{i=1}^{n} q_i |\psi_i\rangle\langle\psi_i|. \tag{1.3}$$

One immediate benefit of this representation by the density operator is that the marginal state that was discussed in Sect. 1.2.2 is represented by a unique density operator, i.e.,

$$\hat{\rho}_A = \sum_{j=1}^{d'} p_j |\phi_j\rangle_{AA}\langle\phi_j| = \sum_{j=1}^{d'} p'_j |\phi'_j\rangle_{AA}\langle\phi'_j| = \mathrm{Tr}_B |\Psi\rangle_{ABAB}\langle\Psi| \tag{1.4}$$

that holds under Eq. (1.2). This operator is called the *marginal* density operator of system $A$ for the whole state $|\Psi\rangle_{AB}$.

Because any positive operator $\hat{\rho}$ with a unit trace can be written in a *diagonal* form $\hat{\rho} = \sum_i \lambda_i |u_i\rangle\langle u_i|$ using nonnegative eigenvalues $\{\lambda_i\}$ with $\sum_i \lambda_i = 1$ and orthonormal eigenvectors $\{|u_i\rangle\}$, $\hat{\rho}$ is the density operator for an ensemble $\{(\lambda_i, |u_i\rangle)\}_i$. Therefore, any density operator is associated with at least one physical state.

When an orthogonal measurement with a basis $\{|u_j\rangle\}_j$ is performed on a mixed state $\{(q_i, |\psi_i\rangle)\}_i$, the probability of the outcome $j$ is calculated using Rule 3 to be $p_j = \sum_i q_i|\langle u_j|\psi_i\rangle|^2 = \langle u_j|\hat{\rho}|u_j\rangle$. This shows that the statistics of the measurement outcome depend only on the density operator. This also shows that each physical state is associated with a single density operator. Consider two mixed states with different density operators $\hat{\rho}$ and $\hat{\rho}'(\neq \hat{\rho})$. Because $|u\rangle \in \mathscr{H}$ exists with $\langle u|(\hat{\rho} - \hat{\rho}')|u\rangle \neq 0$, a measurement leading to different statistics between the two states also exists. The two states are therefore distinct. This fact implies that the density operator can be determined using a map from the set of physical states. As shown earlier, this map is surjective.

The remaining question is whether this map is bijective. At this point, it might not be injective, i.e., different mixed states could be associated with the same density operator. We will provide the answer to this question in Sect. 1.2.5, after we discuss the important properties of bipartite pure states in Sect. 1.2.4.

### *1.2.4 Properties of Bipartite Pure States*

First, we consider how a general bipartite pure state $|\Psi\rangle_{AB}$ can be written in terms of the orthonormal bases $\{|u_i\rangle_A\}_i$ and $\{|v_j\rangle_B\}_j$ for the subsystems $A$ and $B$. Because $\{|u_i\rangle_A|v_j\rangle_B\}_{i,j}$ is a basis of $\mathscr{H}_{AB}$, it is always possible to decompose $|\Psi\rangle_{AB}$ as $|\Psi\rangle_{AB} = \sum_{i,j} c_{i,j}|u_i\rangle_A|v_j\rangle_B$. The special aspect of bipartite states is that a much simpler form of decomposition, $|\Psi\rangle_{AB} = \sum_i c_i|u_i\rangle_A|v_i\rangle_B$, is available if we select $\{|u_i\rangle_A\}_i$ and $\{|v_j\rangle_B\}_j$ appropriately for the given vector $|\Psi\rangle_{AB}$. This decomposition is called *Schmidt decomposition*, and it will be convenient to describe Schmidt decomposition in the form of the following theorem.

**Theorem 2.** *Let $|\Psi\rangle_{AB} \in \mathscr{H}_{AB} = \mathscr{H}_A \otimes \mathscr{H}_B$ be a normalized vector that represents a pure state of a bipartite system AB. Let $\hat{\rho}_A = \mathrm{Tr}_B|\Psi\rangle_{AB\,AB}\langle\Psi|$ be the marginal density operator of system A, and let s be the rank of $\hat{\rho}_A$. For any orthonormal set of vectors $\{|u_i\rangle_A\}_{i=1,\ldots,s} \subset \mathscr{H}_A$ that diagonalizes $\hat{\rho}_A$ as $\hat{\rho}_A = \sum_{i=1}^s p_i|u_i\rangle_{A\,A}\langle u_i|$ with $p_i > 0 (i = 1, \ldots, s)$, there is an orthonormal set of vectors $\{|v_i\rangle_B\}_{i=1,\ldots,s} \subset \mathscr{H}_B$, such that*

$$|\Psi\rangle_{AB} = \sum_{i=1}^{s} \sqrt{p_i}|u_i\rangle_A|v_i\rangle_B. \tag{1.5}$$

*Proof.* Define the unnormalized vectors $|\tilde{v}_i\rangle_B := {}_A\langle u_i||\Psi\rangle_{AB}$. We then have $|\Psi\rangle_{AB} = \sum_{i=1}^s |u_i\rangle_A|\tilde{v}_i\rangle_B$. We see that ${}_B\langle\tilde{v}_i|\tilde{v}_j\rangle_B = \mathrm{Tr}|\tilde{v}_j\rangle_{BB}\langle\tilde{v}_i| = {}_A\langle u_j|\mathrm{Tr}_B(|\Psi\rangle_{ABAB}\langle\Psi|)|u_i\rangle_A = {}_A\langle u_j|\hat{\rho}_A|u_i\rangle_A = p_i\delta_{i,j}$, where $\delta_{i,j} = 1$ if $i = j$, and otherwise $\delta_{i,j} = 0$. Thus, if we define $|v_i\rangle_B := |\tilde{v}_i\rangle_B/\sqrt{p_i}$, $\{|v_i\rangle_B\}_i$ is an orthonormal set that satisfies Eq. (1.5). □

The number $s$ is often called the *Schmidt number* of the state $|\Psi\rangle_{AB}$. If $s$ is smaller than $\dim \mathscr{H}_A$ or $\dim \mathscr{H}_B$, we can always augment the orthonormal sets to form orthonormal bases.

Next, we introduce a concept that is opposite to the concept of the marginal density operator for a bipartite pure state. For a given density operator $\hat{\rho}_A$ of subsystem $A$, a *purification* of the density operator is defined to be a pure state $|\Phi\rangle_{AB}$ of the composite system $AB$ that satisfies $\mathrm{Tr}_B|\Phi\rangle_{AB\,AB}\langle\Phi| = \hat{\rho}_A$. In contrast to the marginal density operator, which is unique to a given state $|\Psi\rangle_{AB}$, the purification of a given density operator $\hat{\rho}_A$ is not unique and there are many bipartite pure states that can be regarded as purifications of $\hat{\rho}_A$. However, they are connected by a simple relation [1, 2] that is given as follows.

**Theorem 3.** *For any two purifications $|\Phi\rangle_{AB}$, $|\Phi'\rangle_{AB} \in \mathscr{H}_{AB} = \mathscr{H}_A \otimes \mathscr{H}_B$ of the same density operator $\hat{\rho}_A$, there is a unitary operator $\hat{V}_B : \mathscr{H}_B \to \mathscr{H}_B$ such that*

$$|\Phi'\rangle_{AB} = (\hat{1}_A \otimes \hat{V}_B)|\Phi\rangle_{AB}. \tag{1.6}$$

*Proof.* When we write down a diagonal form $\hat{\rho}_A = \sum_{i=1}^{s} p_i|u_i\rangle_{AA}\langle u_i|$, Theorem 2 ensures that the purifications are decomposed as $|\Phi\rangle_{AB} = \sum_{i=1}^{s} \sqrt{p_i}|u_i\rangle_A|v_i\rangle_B$ and $|\Phi'\rangle_{AB} = \sum_{i=1}^{s} \sqrt{p_i}|u_i\rangle_A|v_i'\rangle_B$. Because $\{|v_i\rangle_B\}_i$ and $\{|v_i'\rangle_B\}_i$ are orthonormal sets, a unitary operator $\hat{V}_B$ exists such that $|v_i'\rangle_B = \hat{V}_B|v_i\rangle_B$ for all $i$. □

This theorem is quite simple but has deeper consequences. Suppose that Alice holds system $A$ and Bob holds system $B$, and assume that only Bob knows whether the system $AB$ is in state $|\Phi\rangle_{AB}$ or in state $|\Phi'\rangle_{AB}$. There are then only two possible situations: (i) The marginal density operators of subsystem $A$ are different for $|\Phi\rangle_{AB}$ and $|\Phi'\rangle_{AB}$, and thus Alice can locally distinguish state $|\Phi\rangle_{AB}$ from state $|\Phi'\rangle_{AB}$ to some extent. (ii) The marginal density operators of subsystem $A$ are the same and according to Theorem 3, Bob can switch locally between state $|\Phi\rangle_{AB}$ and state $|\Phi'\rangle_{AB}$. As a result, we see that there is no situation whatsoever in which Alice is unable to distinguish between the two states locally *and* Bob is unable to switch between the states locally. This property has led to the no-go theorem for unconditionally secure bit commitment [3, 4].

### 1.2.5 Physical States and Density Operators

We are now in a position to prove that there is a one-to-one correspondence between the physical states and the density operators. Consider two states represented by the ensembles $\{(p_j, |\phi_j\rangle_A)\}_{j=1,\ldots,d}$ and $\{(p_j', |\phi_j'\rangle_A)\}_{j=1,\ldots,d'}$, which are associated with the same density operator $\hat{\rho}_A$. We will show that these two states are in fact the same state [1, 2].

Without loss of generality, we may assume that $d \le d'$. If $d < d'$, we can augment the ensemble $\{(p_j, |\phi_j\rangle_A)\}_{j=1,\ldots,d}$ in an equivalent manner to $\{(p_j, |\phi_j\rangle_A)\}_{j=1,\ldots,d'}$

by adding dummy states $|\phi_j\rangle_A$ with $p_j = 0$. Consider another system $B$ with a Hilbert space $\mathscr{H}_B$ with dimension $d'$, and take an orthonormal basis $\{|v_j\rangle_B\}_{j=1,\dots,d'}$. We then define the bipartite states $|\Psi\rangle_{AB} := \sum_{j=1}^{d'} \sqrt{p_j}|\phi_j\rangle_A|v_j\rangle_B$ and $|\Psi'\rangle_{AB} := \sum_{j=1}^{d'} \sqrt{p'_j}|\phi'_j\rangle_A|v_j\rangle_B$, which both have $\hat{\rho}_A$ as their marginal density operator. From Theorem 3, there is a unitary operator $\hat{V}_B$ with $|\Psi'\rangle_{AB} = (\hat{1}_A \otimes \hat{V}_B)|\Psi\rangle_{AB}$. We define another orthonormal basis $\{|v'_j\rangle_B\}_{j=1,\dots,d'}$ using $|v'_j\rangle := \hat{V}_B^\dagger|v_j\rangle$. It is then simple to confirm that the requisite of Lemma 1, Eq. (1.2), holds, and thus the two states are the same state. When combined with the previous observation in Sect. 1.2.3, we can conclude that:

> There is a one-to-one correspondence between the set of physical states and the set of density operators.

Having established that the density operators are conceptually an ideal description of the physical states, it is natural to expect that the basic and derived rules will be equally well stated when using the density operators in place of vectors to represent the physical states. In fact, by carefully following the definition, we obtain the following list of formulas.

| | | |
|---|---|---|
| Unitary transformation | $\|\phi_{\text{out}}\rangle = \hat{U}\|\phi_{\text{in}}\rangle$ | $\hat{\rho}_{\text{out}} = \hat{U}\hat{\rho}_{\text{in}}\hat{U}^\dagger$ |
| Orthogonal measurement | $p_j = \|\langle u_j\|\phi_{\text{in}}\rangle\|^2$ | $p_j = \langle u_j\|\hat{\rho}_{\text{in}}\|u_j\rangle$ |
| Local preparation | $\|\phi\rangle_A \otimes \|\psi\rangle_B$ | $\hat{\rho}_A \otimes \hat{\rho}_B$ |
| Measurement on subsystem | $\sqrt{p_j}\|\phi_j\rangle_A = {}_B\langle v_j\|\|\Psi\rangle_{AB}$ | $p_j\hat{\rho}_A^{(j)} = {}_B\langle v_j\|\hat{\rho}_{AB}\|v_j\rangle_B$ |
| Preparation by mixing | $\hat{\rho} = \sum_i q_i\|\phi_i\rangle\langle\phi_i\|$ | $\hat{\rho} = \sum_i q_i\hat{\rho}^{(i)}$ |
| Marginal state | $\hat{\rho}_A = \text{Tr}_B\|\Psi\rangle_{AB\,AB}\langle\Psi\|$ | $\hat{\rho}_A = \text{Tr}_B\hat{\rho}_{AB}$ |

Distinction is made between the pure and mixed states based simply on the rank of the density operator. The state is pure if and only if the rank of its density operator $\hat{\rho}$ is 1, in which case it can be written as $\hat{\rho} = |\phi\rangle\langle\phi|$ using the normalized vector $|\phi\rangle$. The opposite extreme may be the case of the operators with maximal rank, which is equal to the dimension $d$ of the Hilbert space. Among these operators, the state where $\hat{\rho} = \hat{1}/d$ has the unique property of invariance under all unitary transformations, and is called the *maximally mixed* state.

Classification of the density operator can be related to the classification of the bipartite pure states through purification. The Schmidt number of a specific purification is equal to the rank of the density operator. The purification of a rank-one density operator, $\hat{\rho}_A = |u\rangle_{A\,A}\langle u|$, is a product state in the form of $|u\rangle_A|v\rangle_B$, while the purification of a nonpure density operator is an entangled state. The purification of a maximally mixed state is called a *maximally entangled* state. Under Schmidt decomposition of Eq. (1.5), a maximally entangled state $|\Phi\rangle_{AB}$ is written as

$$|\Phi\rangle_{AB} = \frac{1}{\sqrt{d}}\sum_{i=1}^{d}|u_i\rangle_A|v_i\rangle_B, \tag{1.7}$$

where $d$ is the dimension of $\mathscr{H}_A$.

## 1.3 Qubits

The simplest of the physical systems is a two-level system that is associated with a Hilbert space of dimension 2, and is called a qubit. For a qubit, the general states, the orthogonal measurements, and the unitary transformations can be conveniently visualized using a three-dimensional image called the Bloch representation.

### *1.3.1 Pauli Operators*

Consider a qubit and choose an orthonormal basis $\{|0\rangle, |1\rangle\}$ of its Hilbert space $\mathscr{H}$ as the standard basis. We define a set of three operators, called *Pauli operators*, as $\hat{\sigma}_x = \hat{\sigma}_1 := |0\rangle\langle 1| + |1\rangle\langle 0|$, $\hat{\sigma}_y = \hat{\sigma}_2 := -i|0\rangle\langle 1| + i|1\rangle\langle 0|$, and $\hat{\sigma}_z = \hat{\sigma}_3 := |0\rangle\langle 0| - |1\rangle\langle 1|$. In the matrix representation under the standard basis, they are written as

$$\hat{\sigma}_x = \hat{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \hat{\sigma}_y = \hat{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \hat{\sigma}_z = \hat{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.8}$$

They satisfy the following commutation and anti-commutation relations:

$$[\hat{\sigma}_i, \hat{\sigma}_j] = 2i\epsilon_{ijk}\hat{\sigma}_k \ \text{ and } \ \{\hat{\sigma}_i, \hat{\sigma}_j\} = 2\delta_{i,j}\hat{1}, \tag{1.9}$$

where $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$, $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$. The Levi-Civita symbol $\epsilon_{ijk}$ is zero, except for $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$ and $\epsilon_{321} = \epsilon_{132} = \epsilon_{213} = -1$, and the Einstein notation is used to omit the summation.

Together with $\hat{\sigma}_0 := \hat{1}$, we have four self-adjoint and unitary operators. These satisfy the orthogonality relations,

$$\mathrm{Tr}(\hat{\sigma}_\mu \hat{\sigma}_\nu) = 2\delta_{\mu,\nu} \tag{1.10}$$

for $\mu, \nu = 0, 1, 2, 3$. Every linear operator $\hat{A}$ acting on $\mathscr{H}$ is uniquely decomposed as $\hat{A} = (P_0\hat{1} + P_x\hat{\sigma}_x + P_y\hat{\sigma}_y + P_z\hat{\sigma}_z)/2$, where the four complex parameters $(P_0, P_x, P_y, P_z)$ can be determined using $P_0 = \mathrm{Tr}(\hat{A})$, $P_x = \mathrm{Tr}(\hat{\sigma}_x\hat{A})$, $P_y = \mathrm{Tr}(\hat{\sigma}_y\hat{A})$, and $P_z = \mathrm{Tr}(\hat{\sigma}_z\hat{A})$. It is convenient to regard $\boldsymbol{P} := (P_x, P_y, P_z)$ as a three-dimensional vector, and to define $\hat{\boldsymbol{\sigma}} := (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$ as well. We denote the inner product between these vectors as $\boldsymbol{P} \cdot \hat{\boldsymbol{\sigma}} := P_x\hat{\sigma}_x + P_y\hat{\sigma}_y + P_z\hat{\sigma}_z$, and the squared norm as $|\boldsymbol{P}|^2 := P_x^2 + P_y^2 + P_z^2$. Using the vector notation, we have

$$\hat{A} = (P_0\hat{1} + \boldsymbol{P} \cdot \hat{\boldsymbol{\sigma}})/2 \tag{1.11}$$

with $P_0 = \mathrm{Tr}(\hat{A})$ and $\boldsymbol{P} = \mathrm{Tr}(\hat{\boldsymbol{\sigma}}\hat{A})$.

Because $\hat{A}^\dagger = (\bar{P}_0\hat{1} + \bar{\boldsymbol{P}} \cdot \hat{\boldsymbol{\sigma}})/2$, $\hat{A}$ is self-adjoint if and only if both $P_0$ and $\boldsymbol{P}$ are real. For a self-adjoint operator $\hat{A}$, it is simple to show that $\det(\hat{A}) = (P_0^2 - |\boldsymbol{P}|^2)/4$, and that the two eigenvalues of $\hat{A}$ are $(P_0 \pm |\boldsymbol{P}|)/2$. Therefore, $\hat{A}$ is positive if and only if $\boldsymbol{P}$ is real and $P_0 \geq |\boldsymbol{P}|$.

### 1.3.2 General States of a Qubit

Because a density operator $\hat{\rho}$ is positive and has a unit trace, application of the decomposition of Eq. (1.11) leads to

$$\hat{\rho} = (\hat{1} + \boldsymbol{P} \cdot \hat{\boldsymbol{\sigma}})/2 \tag{1.12}$$

where the real vector $\boldsymbol{P} = \mathrm{Tr}(\hat{\boldsymbol{\sigma}}\hat{\rho})$ satisfies $|\boldsymbol{P}| \leq 1$. We see that the density operators, and thus the general states of a qubit, are uniquely represented by three-dimensional real vectors $\boldsymbol{P} = (P_x, P_y, P_z)$ with lengths no greater than unity. These vectors are called the *Bloch vectors*, and representation of the qubit states using these Bloch vectors is called *Bloch representation*. As shown in Fig. 1.1a, a Bloch vector is visualized in an *xyz*-Cartesian coordinate system as an arrow stemming from the origin and reaching a point $(P_x, P_y, P_z)$ on or inside of a sphere of unit radius, which is called a *Bloch sphere*.

As shown in Sect. 1.2.5, the rank of $\hat{\rho}$ is 1 when it is a pure state, and for a qubit this implies that the smaller of the eigenvalues of $\hat{\rho}$, $(1-|\boldsymbol{P}|)/2$, is zero. A pure state is thus represented by a Bloch vector of length $|\boldsymbol{P}| = 1$, with the vector tip reaching the Bloch sphere. For a mixed (and nonpure) state, the length of the Bloch vector is shorter ($|\boldsymbol{P}| < 1$). The maximally mixed state with $\hat{\rho} = \hat{1}/2$ is represented by the zero vector $\boldsymbol{P} = \boldsymbol{0}$.

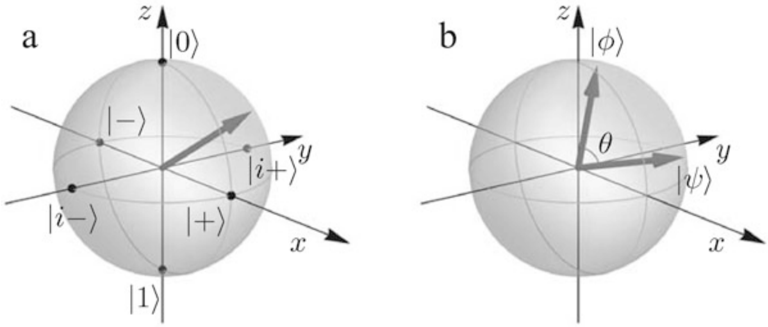

**Fig. 1.1** (**a**) Bloch sphere and a Bloch vector. The six pure states on one of the three axes, where $|\pm\rangle := (|0\rangle \pm |1\rangle)/\sqrt{2}$ and $|i\pm\rangle := (|0\rangle \pm i|1\rangle)/\sqrt{2}$, are also shown. (**b**) A pair of pure states $|\phi\rangle$ and $|\psi\rangle$, with $|\langle\phi|\psi\rangle| = \cos(\theta/2)$

Bloch vectors should not be confused with the vectors of the Hilbert space. Bloch vectors belong to a three-dimensional real vector space, while the Hilbert space of a qubit is a complex two-dimensional vector space. Consider two pure states, $\hat{\rho}_\phi = |\phi\rangle\langle\phi|$ and $\hat{\rho}_\psi = |\psi\rangle\langle\psi|$, with Bloch vectors $\boldsymbol{P}_\phi$ and $\boldsymbol{P}_\psi$, respectively. When $\boldsymbol{P}_\phi \cdot \boldsymbol{P}_\psi = \cos\theta$, then the angle between the two Bloch vectors is $\theta$ (see Fig. 1.1b). In contrast, based on Eq. (1.10), we have $|\langle\phi|\psi\rangle|^2 = \mathrm{Tr}(\hat{\rho}_\phi\hat{\rho}_\psi) = (1+\boldsymbol{P}_\phi\cdot\boldsymbol{P}_\psi)/2 = \cos^2(\theta/2)$, which implies that the angle between the two vectors of the Hilbert space is $\theta/2$. For two orthogonal pure states, $\theta/2 = \pi/2$ implies that the corresponding pair of Bloch vectors point in opposite directions.

### 1.3.3 Orthogonal Measurement on a Qubit

Let us interpret an orthogonal measurement using the basis $\{|u_0\rangle, |u_1\rangle\}$ in terms of Bloch representation. We define the Bloch vectors $\boldsymbol{P}_0$ and $\boldsymbol{P}_1$ for the basis states using $\hat{\rho}_j := |u_j\rangle\langle u_j| = (\hat{1} + \boldsymbol{P}_j \cdot \hat{\boldsymbol{\sigma}})/2$. Because the orthogonality $\langle u_0|u_1\rangle = 0$ implies that $\boldsymbol{P}_1 = -\boldsymbol{P}_0$, the orthogonal measurement is completely characterized by the unit vector $\boldsymbol{P}_0$, which is a direction in the three-dimensional space.

Suppose that the measured qubit is initially in the state given by $\hat{\rho} = (\hat{1}+\boldsymbol{P}\cdot\hat{\boldsymbol{\sigma}})/2$. The probabilities of outcome $j = 0, 1$ are then calculated to be $p_j = \langle u_j|\hat{\rho}|u_j\rangle = \mathrm{Tr}(\hat{\rho}_j\hat{\rho}) = (1 + \boldsymbol{P}_j \cdot \boldsymbol{P})/2$, leading to

$$p_0 = (1 + \boldsymbol{P}_0 \cdot \boldsymbol{P})/2 \;\text{ and }\; p_1 = (1 - \boldsymbol{P}_0 \cdot \boldsymbol{P})/2. \tag{1.13}$$

This shows that the probabilities are essentially determined by projection of the measured Bloch vector $\boldsymbol{P}$ along the direction $\boldsymbol{P}_0$ that was specified by the measurement, with appropriate scaling (see Fig. 1.2a).

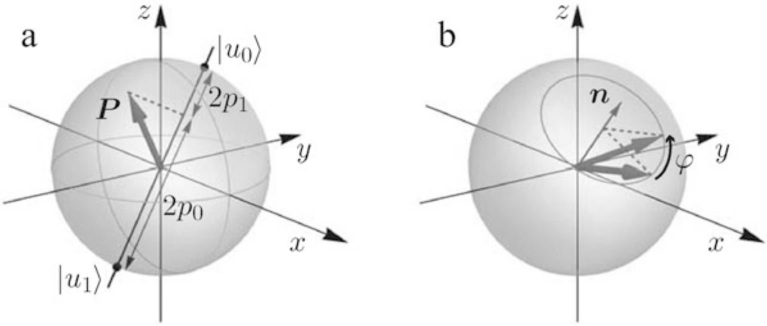

**Fig. 1.2** (**a**) Orthogonal measurement with a basis $\{|u_j\rangle\}_{j=0,1}$. The Bloch vector $\boldsymbol{P}$ of the input state determines the probability $p_j$ of the outcome $j$. (**b**) The Bloch vector rotates in a unitary transformation with $\hat{U}(\boldsymbol{n}, \varphi)$

### 1.3.4 Unitary Transformation on a Qubit

We now discuss how the Bloch vector of a physical state changes under a unitary transformation. We limit ourselves to the unitary operators $\hat{U}$ that belong to a set called $SU(2)$, and are characterized by the condition $\det \hat{U} = 1$. This does not lose generality because $\hat{U}(\theta) := e^{i\theta/2}\hat{U}$ for real $\theta$ transforms the state $\hat{\rho}$ to $\hat{U}(\theta)\hat{\rho}\hat{U}(\theta)^\dagger = \hat{U}\hat{\rho}\hat{U}^\dagger$, which is independent of $\theta$. All $\hat{U}(\theta)$ physically represent the same transformation, and we are thus allowed to choose one that satisfies $\det \hat{U}(\theta) = e^{i\theta} \det \hat{U} = 1$ and thus $\hat{U}(\theta) \in SU(2)$. Note that the correspondence is not one-to-one but in fact two-to-one, because $-\hat{U}(\theta) = \hat{U}(\theta + 2\pi)$ also belongs to $SU(2)$.

The elements of $SU(2)$ are conveniently parametrized as follows. Any $\hat{U} \in SU(2)$ can be written in the diagonal form $\hat{U} = e^{-i\varphi/2}|u_0\rangle\langle u_0| + e^{i\varphi/2}|u_1\rangle\langle u_1|$ with $\langle u_0|u_1\rangle = 0$. We may then write $\hat{U} = \exp(-i\varphi\hat{S}/2)$ with $\hat{S} := |u_0\rangle\langle u_0| - |u_1\rangle\langle u_1|$, which is self-adjoint, traceless, and has eigenvalues of $\pm 1$. Using the decomposition of Eq. (1.11), we find that $\hat{S}$ is written as $\hat{S} = \boldsymbol{P}\cdot\hat{\boldsymbol{\sigma}}/2$ with $|\boldsymbol{P}| = 2$. By introducing a unit vector $\boldsymbol{n} := \boldsymbol{P}/2$, we conclude that the elements of $SU(2)$ can be parametrized as

$$\hat{U}(\boldsymbol{n}, \varphi) := \exp[-i(\varphi/2)\boldsymbol{n}\cdot\hat{\boldsymbol{\sigma}}]. \tag{1.14}$$

We are interested in how the Bloch vector evolves when the density operator evolves under a unitary transformation. Noting that $\hat{U}(\boldsymbol{n}, \varphi+\varphi') = \hat{U}(\boldsymbol{n}, \varphi')\hat{U}(\boldsymbol{n}, \varphi)$ holds in general, we see that it is sufficient to focus on the transformations given by $\hat{U}(\boldsymbol{n}, \delta\varphi)$, where $\delta\varphi$ is infinitesimally small. A general transformation $\hat{U}(\boldsymbol{n}, \varphi)$ is then understood as a result of sequential application of these infinitesimal transformations.

Under the transformation $\hat{U}(\boldsymbol{n}, \delta\varphi)$, a Bloch vector $\boldsymbol{P} := \mathrm{Tr}(\hat{\boldsymbol{\sigma}}\hat{\rho})$ evolves into $\boldsymbol{P} + \delta\boldsymbol{P} = \mathrm{Tr}(\hat{\boldsymbol{\sigma}}\hat{\rho}')$ with $\hat{\rho}' := \hat{U}(\boldsymbol{n}, \delta\varphi)\hat{\rho}\hat{U}(\boldsymbol{n}, \delta\varphi)^\dagger$. Using $\hat{U}(\boldsymbol{n}, \delta\varphi) \cong \hat{1} - i(\delta\varphi/2)\boldsymbol{n}\cdot\hat{\boldsymbol{\sigma}}$ and collecting the terms up to the first order in $\delta\varphi$, we find that $\delta\boldsymbol{P} = \mathrm{Tr}(\hat{\boldsymbol{\sigma}}\hat{\rho}') - \mathrm{Tr}(\hat{\boldsymbol{\sigma}}\hat{\rho}) = -i(\delta\varphi/2)\mathrm{Tr}([\hat{\boldsymbol{\sigma}}, \boldsymbol{n}\cdot\hat{\boldsymbol{\sigma}}]\hat{\rho})$. From Eq. (1.9), we obtain $[\hat{\sigma}_i, n_j\hat{\sigma}_j] = 2i\epsilon_{ijk}n_j\hat{\sigma}_k$ under the Einstein notation, which implies that $[\hat{\boldsymbol{\sigma}}, \boldsymbol{n}\cdot\hat{\boldsymbol{\sigma}}] = 2i\boldsymbol{n}\times\hat{\boldsymbol{\sigma}}$. Therefore, $\hat{U}(\boldsymbol{n}, \delta\varphi)$ induces an infinitesimal change in the Bloch vector, which is given by

$$\delta\boldsymbol{P} = \delta\varphi\boldsymbol{n}\times\boldsymbol{P}. \tag{1.15}$$

This is equal to the infinitesimal change in rotation around axis $\boldsymbol{n}$ by the angle $\delta\varphi$. We thus conclude that the Bloch vectors rotate around axis $\boldsymbol{n}$ by angle $\varphi$ under the general unitary transformation $\hat{U}(\boldsymbol{n}, \varphi)$ (see Fig. 1.2b). Notable examples include the $Z$ gate with $\hat{U}((0, 0, 1), \pm\pi) = \mp i\hat{\sigma}_z$, the $X$ gate with $\hat{U}((1, 0, 0), \pm\pi) = \mp i\hat{\sigma}_x$, and the Hadamard gate with $\hat{U}((2^{-1/2}, 0, 2^{-1/2}), \pm\pi) = \mp 2^{-1/2}i(\hat{\sigma}_z + \hat{\sigma}_x)$.

## 1.4 Generalized Measurements and Quantum Operations

The basic set of rules that we adopted in Sect. 1.1 dictated that we can carry out unitary transformations and orthogonal measurements on a physical system (Rules 2 and 3). Here, we extend the repertoire of what we can do to a physical system by using an auxiliary system as a workspace. We also clarify how far this extension goes, and draw a clear line between what we can and cannot do.

### *1.4.1 Use of Auxiliary Systems*

Suppose that we want to operate on a physical system $A$. Let $\hat{\rho}_{\text{in}}$ be the density operator for the initial state of the system $A$. We first prepare an auxiliary system $E$, which has a Hilbert space $\mathscr{H}_E$ of dimension $s$, in a fixed pure state $|\phi_{\text{ini}}\rangle_E$. We then let the systems $A$ and $E$ interact with each other such that the unitary transformation described by the unitary operator $\hat{U}_{AE} : \mathscr{H}_A \otimes \mathscr{H}_E \to \mathscr{H}_A \otimes \mathscr{H}_E$ occurs. Finally, we perform an orthogonal measurement on system $E$ with an orthonormal basis $\{|j\rangle_E\}_{j=1,\dots,s}$ of $\mathscr{H}_E$. The output of the operation is the classical variable $j$ and the final quantum state $\hat{\rho}_{\text{out}}^{(j)}$ of system $A$, which may depend on the value of $j$. This can thus be regarded as conducting a state transformation and performing a measurement at the same time. Using the rules that were summarized in Sect. 1.2.5, we can easily show how the final state $\hat{\rho}_{\text{out}}^{(j)}$ and the probability $p_j$ of obtaining $j$ are related to the initial state:

$$p_j \hat{\rho}_{\text{out}}^{(j)} = {}_E\langle j|\hat{U}_{AE}(\hat{\rho}_{\text{in}} \otimes |\phi_{\text{ini}}\rangle_{EE}\langle\phi_{\text{ini}}|)\hat{U}_{AE}^{\dagger}|j\rangle_E. \tag{1.16}$$

We sometimes encounter a situation where the input and the output are different physical systems. For example, in the photoelectric effect, light is incident on a metal but an electron comes out of the metal. In such a case, we would regard the light field as the input system $A$, and the metal, including the electron that is eventually emitted, as the auxiliary system $E$. The whole system is the composite of $A$ and $E$. The output system, i.e., the electron, is a subsystem of the composite system $AE$, and we call it system $A'$. The rest of system $AE$ is then called system $E'$. In short, we have introduced two different ways to decompose the entire system into two subsystems, $AE$ and $A'E'$. Mathematically, this corresponds to an equivalence relation $\mathscr{H}_A \otimes \mathscr{H}_E = \mathscr{H}_{A'} \otimes \mathscr{H}_{E'}$.

We can now generalize the strategy for use of an auxiliary system to include cases where the output system is not necessarily the same as the input system, as shown in Fig. 1.3. It is convenient to regard the unitary operator $\hat{U}_{AE}$ as a linear map $\hat{U} : \mathscr{H}_A \otimes \mathscr{H}_E \to \mathscr{H}_{A'} \otimes \mathscr{H}_{E'}$, where we dropped the subscript $AE$. Let $s'$ be the dimension of $\mathscr{H}_{E'}$. The orthogonal measurement is performed on system $E'$ with an orthonormal basis $\{|j\rangle_{E'}\}_{j=1,\dots,s'}$. Equation (1.16) is then generalized as

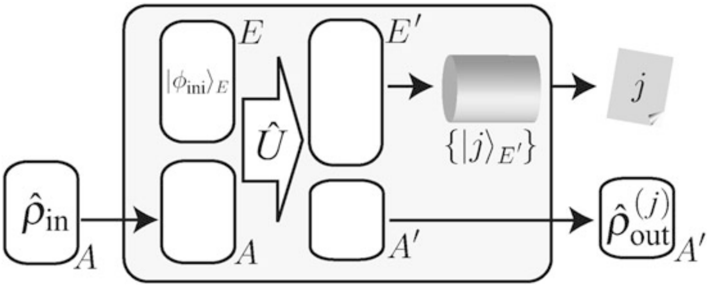
$E$
$|\phi_{\mathrm{ini}}\rangle_E$
$\hat{U}$
$E'$
$j$
$\{|j\rangle_{E'}\}$
$\hat{\rho}_{\mathrm{in}}$
$A$
$A$
$A'$
$\hat{\rho}_{\mathrm{out}}^{(j)}$
$A'$

**Fig. 1.3** Use of an auxiliary system in operation on the physical system $A$. An auxiliary system $E$ is prepared in a fixed pure state $|\phi_{\mathrm{ini}}\rangle_E$, and the unitary transformation $\hat{U}$ is applied to systems $A$ and $E$. System $A'$, which is part of the whole system $AE$, is released as an output. The remaining system, $E'$, is measured to produce the outcome $j$

$$p_j\hat{\rho}_{\mathrm{out}}^{(j)} = {}_{E'}\langle j|\hat{U}(\hat{\rho}_{\mathrm{in}} \otimes |\phi_{\mathrm{ini}}\rangle_{EE}\langle\phi_{\mathrm{ini}}|)\hat{U}^\dagger|j\rangle_{E'}. \tag{1.17}$$

It is convenient to introduce the operators $\hat{M}^{(j)} : \mathscr{H}_A \to \mathscr{H}_{A'}$, which are defined by

$$\hat{M}^{(j)} = {}_{E'}\langle j|\hat{U}|\phi_{\mathrm{ini}}\rangle_E. \tag{1.18}$$

Using the relation $\sum_{j=1}^{s'} |j\rangle_{E'E'}\langle j| = \hat{1}_{E'}$, we see that the operators satisfy the normalization condition

$$\sum_{j=1}^{s'} \hat{M}^{(j)\dagger}\hat{M}^{(j)} = \hat{1}_A. \tag{1.19}$$

The set of operators $\{\hat{M}^{(j)} : \mathscr{H}_A \to \mathscr{H}_{A'}\}$ that satisfies the above relationship are often called *Kraus* operators. Using these operators, Eq. (1.17) can be simplified as

$$p_j\hat{\rho}_{\mathrm{out}}^{(j)} = \hat{M}^{(j)}\hat{\rho}_{\mathrm{in}}\hat{M}^{(j)\dagger}, \tag{1.20}$$

where the input-output relationship is stated without any reference to the auxiliary systems $E$ and $E'$.

In the above argument, we started with a given operator $\hat{U}$ that represented a unitary transformation of the composite system to determine the Kraus operators for the simplified relationship of Eq. (1.20). As we will see, this process can be reversed, i.e., for any given set of Kraus operators $\{\hat{M}^{(j)}\}$ that satisfies Eq. (1.19), there is[3] a unitary operator $\hat{U}$ that satisfies Eq. (1.18). Let $\{|u_i\rangle_A\}_{i=1,\ldots,d}$ be an orthonormal

[3] Given $s'$, we may choose the dimensions of $\mathscr{H}_E$ and $\mathscr{H}_{E'}$ such that they satisfy $\dim \mathscr{H}_{E'} \geq s'$ and $\dim \mathscr{H}_A \dim \mathscr{H}_E = \dim \mathscr{H}_{A'} \dim \mathscr{H}_{E'}$.

basis of $\mathscr{H}_A$. Then, $\{|u_i\rangle_A \otimes |\phi_{\rm ini}\rangle_E\}_{i=1,\ldots,d}$ is an orthonormal set. We define $|v_i\rangle \in \mathscr{H}_{A'} \otimes \mathscr{H}_{E'}$ by $|v_i\rangle := \sum_{j=1}^{s'} \hat{M}^{(j)}|u_i\rangle \otimes |j\rangle_{E'}$. From Eq. (1.19), it can be shown that $\{|v_i\rangle\}_{i=1,\ldots,d}$ is an orthonormal set. There is thus a unitary operator $\hat{U} : \mathscr{H}_A \otimes \mathscr{H}_E \to \mathscr{H}_{A'} \otimes \mathscr{H}_{E'}$ that connects the two orthonormal sets as $|v_i\rangle = \hat{U}|u_i\rangle_A \otimes |\phi_{\rm ini}\rangle_E$, which leads to Eq. (1.18). We thus conclude that any input-output relationship dictated by the Kraus operators as shown in Eq. (1.20) can be physically implemented by attaching an auxiliary system $E$, applying a suitable unitary transformation over the composite system, and then measuring the subsystem $E'$.

### 1.4.2 Physically Allowed Operations

In Sect. 1.4.1, we extended our ability to operate on physical systems through the rather heuristic use of an auxiliary system. It is natural to expect that the introduction of more complex schemes using two or more auxiliary systems may allow us to further extend the variety of possible operations. Additionally, if we look back on the basic rules in Sect. 1.1, we see that none of the rules require a physical operation to be built up from unitary transformations and orthogonal measurements alone. Nonetheless, we will show here that the input-output relations written in the form of Eq. (1.20) are essentially the only relations that are allowed physically.

Consider a black box that accepts a physical system $A$ as an input, and produces a classical outcome $j = 1, 2, \ldots, s$, while leaving the system $A'$ as an output. Let $d$ be the dimension of $\mathscr{H}_A$. We want to know the way in which the output state $\hat{\rho}_{\rm out}^{(j)}$ and the probability $p_j$ of the outcome are related to a general pure input state $|\phi\rangle_A$. For that purpose, it is convenient to introduce a reference system $B$ with a Hilbert space $\mathscr{H}_B$ of the same dimension $d$. We take the orthonormal bases $\{|i\rangle_A\}_{i=1,\ldots,d}$ and $\{|i\rangle_B\}_{i=1,\ldots,d}$ for $\mathscr{H}_A$ and $\mathscr{H}_B$, respectively, and suppose that the system $AB$ is initially prepared in a maximally entangled state, $|\Phi\rangle_{AB} = d^{-1/2} \sum_{i=1}^{d} |i\rangle_A |i\rangle_B$.

We now explain a frequently used technique called the *relative* states. For any given state $|\phi\rangle_A$, we define the relative state of system $B$, with reference to the maximally entangled state $|\Phi\rangle_{AB}$, as

$$|\phi^*\rangle_B := \sum_{i=1}^{d} |i\rangle_{BA}\langle\phi|i\rangle_A. \tag{1.21}$$

It is then easy to see that

$$d^{-1/2}|\phi\rangle_A = {}_B\langle\phi^*||\Phi\rangle_{AB} \tag{1.22}$$

holds. The definition of the relative state is mutual, i.e., $|\phi^{**}\rangle_A = |\phi\rangle_A$, because $d^{-1/2}|\phi^*\rangle_B = {}_A\langle\phi||\Phi\rangle_{AB}$ also holds.

In light of Theorem 1, this relation has the following meaning. If we conduct an orthogonal measurement on system $B$ with a basis that includes state $|\phi^*\rangle_B$, then the

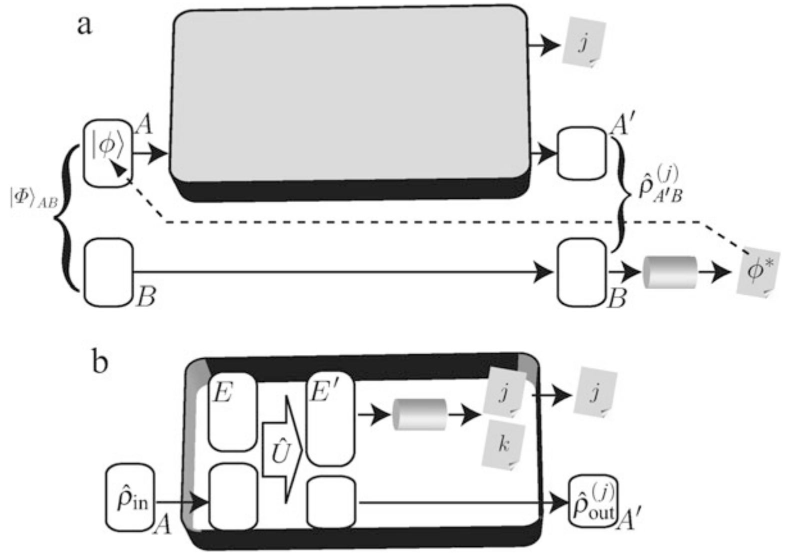

**Fig. 1.4** (**a**) Characterization of a physical operation (the *black box*) by feeding in half of a maximally entangled state. Learning the statistics of the outcome $j$ and the states $\hat{\rho}^{(j)}_{A'B}$ then allows us to fully specify the input-output relationship. (**b**) Looking inside the box. Any physical process is implemented in an equivalent manner with an auxiliary system, a unitary transformation, and an orthogonal measurement

corresponding outcome appears with probability $1/d$, and system $A$ then behaves as if it were initially prepared in state $|\phi\rangle_A$. While this is probabilistic, it offers a type of ex post facto method to prepare system $A$ in the arbitrary state $|\phi\rangle_A$.

We now proceed to the analysis of the black box (see Fig. 1.4a). After preparation of $|\Phi\rangle_{AB}$, suppose that system $A$ is fed to the black box, while system $B$ is left alone. After the black box has produced the outcome $j$, the state of the composite system $A'B$ should be represented by a density operator, which we denote by $\hat{\rho}^{(j)}_{A'B}$. Let $q_j$ be the probability of producing the outcome $j$. Now suppose that we perform an orthogonal measurement with basis $\{|v_i\rangle_B\}$ on system $B$, where $|v_1\rangle_B = |\phi^*\rangle_B$. The outcome $i = 1$ should then appear with probability $r^{(j)}$ and this leaves system $A'$ in state $\hat{\rho}^{(j)}_{A'}$, where

$$r^{(j)}\hat{\rho}^{(j)}_{A'} = {}_B\langle\phi^*|\hat{\rho}^{(j)}_{A'B}|\phi^*\rangle_B. \tag{1.23}$$

However, according to Theorem 1, an event with outcomes $j$ and $i = 1$ must be interpreted as follows. With probability $d^{-1}$, system $A$ is initially prepared in $|\phi\rangle_A$, and is then fed to the black box. This produces outcome $j$ with probability $p_j$, leaving system $A'$ in state $\hat{\rho}^{(j)}_{\text{out}}$. Comparison of the two interpretations leads to $q_j r^{(j)} = d^{-1}p_j$ and $\hat{\rho}^{(j)}_{A'} = \hat{\rho}^{(j)}_{\text{out}}$. Using Eq. (1.23), we then have

$$p_j\hat{\rho}^{(j)}_{\mathrm{out}} = dq_{jB}\langle\phi^*|\hat{\rho}^{(j)}_{A'B}|\phi^*\rangle_B. \tag{1.24}$$

Consider a decomposition of the density operator,

$$\hat{\rho}^{(j)}_{A'B} = \sum_{k=1}^{t^{(j)}} |\tilde{\Psi}^{(j)}_k\rangle_{A'BA'B}\langle\tilde{\Psi}^{(j)}_k|, \tag{1.25}$$

where $|\tilde{\Psi}^{(j)}_k\rangle_{A'B}$ is unnormalized. Noting that ${}_B\langle\phi^*| = \sqrt{d}_{AB}\langle\Phi||\phi\rangle_A$, we see that, for fixed values of $j$ and $k$, the correspondence $|\phi\rangle_A \mapsto \sqrt{dq_j}{}_B\langle\phi^*||\Psi^{(j)}_k\rangle_{A'B}$ is a linear map. Thus, an operator $\hat{M}^{(j,k)} : \mathscr{H}_A \to \mathscr{H}_{A'}$ exists such that

$$\sqrt{dq_j}{}_B\langle\phi^*||\Psi^{(j)}_k\rangle_{A'B} = \hat{M}^{(j,k)}|\phi\rangle_A. \tag{1.26}$$

Equation (1.24) is now written as

$$p_j\hat{\rho}^{(j)}_{\mathrm{out}} = \sum_{k=1}^{t^{(j)}} \hat{M}^{(j,k)}|\phi\rangle_{AA}\langle\phi|\hat{M}^{(j,k)\dagger} \tag{1.27}$$

for the input state $|\phi\rangle_A$. Then, for the general input state $\hat{\rho}_{\mathrm{in}}$ of system $A$, the input-output relationship of the black box is written as

$$p_j\hat{\rho}^{(j)}_{\mathrm{out}} = \sum_{k=1}^{t^{(j)}} \hat{M}^{(j,k)}\hat{\rho}_{\mathrm{in}}\hat{M}^{(j,k)\dagger}. \tag{1.28}$$

Taking the trace of Eq. (1.27) and performing a sum over index $j$, we have $\sum_{j,k}{}_A\langle\phi|\hat{M}^{(j,k)\dagger}\hat{M}^{(j,k)}|\phi\rangle_A = 1$ for arbitrary $|\phi\rangle_A$. Therefore, $\sum_{j,k}\hat{M}^{(j,k)\dagger}\hat{M}^{(j,k)} = \hat{1}_A$ and $\{\hat{M}^{(j,k)}\}$ is a set of Kraus operators.

Equation (1.28) is the most general form of what we can do to a physical system. This equation is merely a trivial extension of Eq. (1.20) in Sect. 1.4.1. Consider a scheme that produces the outcome $(j,k)$ and leaves system $A'$ in state $\hat{\rho}^{(j,k)}_{\mathrm{out}}$, with an input-output relation given by $p_{j,k}\hat{\rho}^{(j,k)}_{\mathrm{out}} = \hat{M}^{(j,k)}\hat{\rho}_{\mathrm{in}}\hat{M}^{(j,k)\dagger}$. As shown[4] in Sect. 1.4.1, this scheme can be implemented by simply attaching an auxiliary system $E$, applying a unitary transformation, and then performing an orthogonal measurement on system $E'$. The original black box is then faithfully simulated using this scheme as shown in Fig. 1.4b, by simply discarding the index $k$ and yielding only the index $j$ as the final outcome.

[4] Regard $(j,k)$ as a single index with the values $1,\ldots,s'$, where $s' = \sum_j t^{(j)}$.

### 1.4.3 Generalized Measurements

By discarding the output quantum state in system $A'$ in the black box that was considered in Sect. 1.4.2, we can obtain the most general form of a physically allowed measurement process, which is called a *generalized measurement*. By taking the trace of Eq. (1.28), we have

$$p_j = \mathrm{Tr}(\hat{F}^{(j)}\hat{\rho}_{\mathrm{in}}), \tag{1.29}$$

where $\hat{F}^{(j)} := \sum_k \hat{M}^{(j,k)\dagger}\hat{M}^{(j,k)}$ is positive and satisfies $\sum_j \hat{F}^{(j)} = \hat{1}_A$. Any measurement must be written in this form.

A set of positive operators $\{\hat{F}^{(j)}\}$ acting on $\mathscr{H}_A$ and satisfying $\sum_j \hat{F}^{(j)} = \hat{1}_A$ is called the *POVM (positive-operator-valued measure)*. For any given POVM $\{\hat{F}^{(j)}\}$, we may define $\hat{M}^{(j)} := (\hat{F}^{(j)})^{1/2}$ and use the argument of Sect. 1.4.1 to construct a generalized measurement that satisfies Eq. (1.29) through the use of an auxiliary system as shown in Fig. 1.3, except that system $A'$ is discarded in this case.

An orthogonal measurement with basis $\{|u_j\rangle_A\}$ is now regarded as a special case of the generalized measurements, when the POVM is chosen to be $\hat{F}^{(j)} = |u_j\rangle_{AA}\langle u_j|$. Note that orthogonal measurements are not necessarily the ideal measurement, and some tasks favor other kinds of generalized measurement. We will provide an example below.

**Unambiguous state discrimination.** Consider a nonorthogonal pair of qubit states, $\{|\phi_0\rangle_A, |\phi_1\rangle_A\}$, with $c := |\langle\phi_0|\phi_1\rangle| > 0$. Suppose that qubit $A$ has been secretly prepared in $|\phi_0\rangle_A$ or in $|\phi_1\rangle_A$ with an equal probability of $q := 1/2$. Consider the strategy used to distinguish between the two states as follows.

Choose $|\phi_j^\perp\rangle_A$ $(j = 0, 1)$ such that ${}_A\langle\phi_j|\phi_j^\perp\rangle_A = 0$ and ${}_A\langle\phi_0^\perp|\phi_1^\perp\rangle_A = c$. Consider a set $\{\hat{F}^{(j)}\}_{j=0,1,2}$ defined by $\hat{F}^{(0)} := (1+c)^{-1}|\phi_1^\perp\rangle_{AA}\langle\phi_1^\perp|$, $\hat{F}^{(1)} := (1+c)^{-1}|\phi_0^\perp\rangle_{AA}\langle\phi_0^\perp|$, and $\hat{F}^{(2)} := \hat{1}_A - \hat{F}^{(0)} - \hat{F}^{(1)}$. Because $|\phi_0^\perp\rangle_A \pm |\phi_1^\perp\rangle_A$ is an eigenvector of $\hat{F}^{(0)} + \hat{F}^{(1)}$ with eigenvalue $(1+c)^{-1}(1 \pm c) \leq 1$, we have $\hat{F}^{(0)} + \hat{F}^{(1)} \leq \hat{1}_A$. Therefore, $\{\hat{F}^{(j)}\}_{j=0,1,2}$ is a POVM, and the corresponding generalized measurement is feasible.

When the outcome of this measurement was $j = 0$, we were certain that the prepared state must be state $|\phi_0\rangle_A$, because $\mathrm{Tr}(\hat{F}^{(0)}|\phi_1\rangle_{AA}\langle\phi_1|) = 0$. Similarly, if the outcome was $j = 1$, the prepared state must be state $|\phi_1\rangle_A$. The overall success probability, i.e., the probability of obtaining $j = 0, 1$ is calculated to be $p_{\mathrm{suc}} := \sum_{j=0,1} q\mathrm{Tr}(\hat{F}^{(j)}|\phi_j\rangle_{AA}\langle\phi_j|) = 1 - c$ [5].

If we are to construct a strategy with a similar lack of ambiguity using orthogonal measurements, we must choose either $\{|\phi_0\rangle_A, |\phi_0^\perp\rangle_A\}$ or $\{|\phi_1\rangle_A, |\phi_1^\perp\rangle_A\}$ as the basis. Regardless of how the two orthogonal measurements are mixed, the success probability is $p_{\mathrm{suc}}^\perp := q|{}_A\langle\phi_0^\perp|\phi_1\rangle_A|^2 = q|{}_A\langle\phi_1^\perp|\phi_0\rangle_A|^2 = (1-c^2)/2$. Thus we see that $p_{\mathrm{suc}} > p_{\mathrm{suc}}^\perp$ for $0 < c < 1$.

### 1.4.4 Quantum Operations

If we discard the outcome $j$ from the black box that was considered in Sect. 1.4.2, then the output density operator of system $A'$ becomes $\hat{\rho}_{\text{out}} := \sum_j p_j \hat{\rho}_{\text{out}}^{(j)} = \sum_{j,k} \hat{M}^{(j,k)} \hat{\rho}_{\text{in}} \hat{M}^{(j,k)\dagger}$. Without loss of generality, we may replace the indices $(j, k)$ with a single index $j$, which results in the general form of the state transformation,

$$\hat{\rho}_{\text{out}} = \sum_j \hat{M}^{(j)} \hat{\rho}_{\text{in}} \hat{M}^{(j)\dagger} \tag{1.30}$$

with $\sum_j \hat{M}^{(j)\dagger} \hat{M}^{(j)} = \hat{1}_A$. Any physical process that takes system $A$ as an input and leaves the same system or another system $A'$ as an output must be written in this form. This type of process is often called a *quantum operation* or a *quantum channel*. Mathematically, the map $\chi : \hat{\rho}_{\text{in}} \mapsto \hat{\rho}_{\text{out}}$ that is written as per Eq. (1.30) is called a *CPTP (completely-positive trace-preserving) map*.

The argument in Sect. 1.4.1 ensures that the right-hand side of Eq. (1.30) can be rewritten as that of Eq. (1.17) summed over $j$, i.e.,

$$\hat{\rho}_{\text{out}} = \text{Tr}_{E'}[\hat{U}(\hat{\rho}_{\text{in}} \otimes |\phi_{\text{ini}}\rangle_{EE}\langle\phi_{\text{ini}}|)\hat{U}^\dagger]. \tag{1.31}$$

Operationally, this simply means that the measurement on system $E'$ shown in Fig. 1.3 is unnecessary. Thus, any quantum channel can be equivalently simulated using a simple three-step process, which consists of preparing the auxiliary system ($E$) in a fixed pure state, applying the unitary transformation, and discarding the subsystem ($E'$). This property is very helpful when it is necessary to prove that some tasks are physically impossible. This type of argument is vital for establishment of an operationally-defined measure of quantum properties, as indicated in the following example.

**Fidelity.** In an experimental demonstration, the quality of the final result is often evaluated in terms of the fidelity $F = \langle\phi_{\text{ideal}}|\hat{\rho}_{\text{exp}}|\phi_{\text{ideal}}\rangle$, where $|\phi_{\text{ideal}}\rangle$ is the desired state and $\hat{\rho}_{\text{exp}}$ is the state that was actually obtained in the experiment. The fidelity $F$ between two general states $\hat{\rho}_1$ and $\hat{\rho}_2$ of system $A$ is defined[5] as the maximum overlap between the purifications of these states in a composite system composed of $A$ and an arbitrary system $R$, i.e.,

$$F(\hat{\rho}_1, \hat{\rho}_2) := \max\{|_{AR}\langle\Psi_1|\Psi_2\rangle_{AR}|^2 : \text{Tr}_R(|\Psi_j\rangle_{ARAR}\langle\Psi_j|) = \hat{\rho}_j, j = 1, 2\}. \tag{1.32}$$

To justify the use of such a quantity in the evaluation of an experiment, we must show that the fidelity $F(\hat{\rho}_1, \hat{\rho}_2)$ is a good measure of the closeness between the two

[5] There is an equivalent method to define the fidelity as $F(\hat{\rho}_1, \hat{\rho}_2) = (\text{Tr}\sqrt{\hat{\rho}_1^{1/2}\hat{\rho}_2\hat{\rho}_1^{1/2}})^2$ [6, 7]. In some of the literature, the quantity $\sqrt{F(\hat{\rho}_1, \hat{\rho}_2)}$ is referred to as the fidelity.

states $\hat{\rho}_1$ and $\hat{\rho}_2$. To enable $F$ to quantify the difficulty in distinguishing between the two states *in principle*, $F$ should not be reduced (and thus the distinguishability should not improve) through the application of any quantum channel $\chi$, i.e.,

$$F(\chi(\hat{\rho}_1), \chi(\hat{\rho}_2)) \geq F(\hat{\rho}_1, \hat{\rho}_2) \tag{1.33}$$

should hold for any CPTP map $\chi$. This can be proved as follows.

Let $|\Phi_j\rangle_{AR}$ be the purifications that achieve the maximum of Eq. (1.32), i.e., $F(\hat{\rho}_1, \hat{\rho}_2) = |{}_{AR}\langle\Phi_1|\Phi_2\rangle_{AR}|^2$. We consider three different cases separately, corresponding to the three steps that are implied in Eq. (1.31).

(i) $\chi(\hat{\rho}_j) = \hat{\rho}_j \otimes |\phi\rangle_{BB}\langle\phi|$. In this case, $|\Psi_j\rangle_{ABR} := |\Phi_j\rangle_{AR}|\phi\rangle_B$ is a purification of $\chi(\hat{\rho}_j)$. Therefore, $F(\chi(\hat{\rho}_1), \chi(\hat{\rho}_2)) \geq |{}_{ABR}\langle\Psi_1|\Psi_2\rangle_{ABR}|^2 = |{}_{AR}\langle\Phi_1|\Phi_2\rangle_{AR}|^2 = F(\hat{\rho}_1, \hat{\rho}_2)$.

(ii) $\chi(\hat{\rho}_j) = \hat{U}_A\hat{\rho}_j\hat{U}_A^\dagger$. In this case, $|\Psi_j\rangle_{AR} := (\hat{U}_A \otimes \hat{1}_R)|\Phi_j\rangle_{AR}$ is a purification of $\chi(\hat{\rho}_j)$. Therefore, $F(\chi(\hat{\rho}_1), \chi(\hat{\rho}_2)) \geq |{}_{AR}\langle\Psi_1|\Psi_2\rangle_{AR}|^2 = |{}_{AR}\langle\Phi_1|\Phi_2\rangle_{AR}|^2 = F(\hat{\rho}_1, \hat{\rho}_2)$.

(iii) $\chi(\hat{\rho}_j) = \mathrm{Tr}_{\tilde{A}}(\hat{\rho}_j)$, where $\tilde{A}$ is a constituent subsystem of system $A$. In this case, $|\Phi_j\rangle_{AR}$ is also regarded as a purification of $\chi(\hat{\rho}_j)$. Therefore, $F(\chi(\hat{\rho}_1), \chi(\hat{\rho}_2)) \geq |{}_{AR}\langle\Phi_1|\Phi_2\rangle_{AR}|^2 = F(\hat{\rho}_1, \hat{\rho}_2)$.

For a general quantum channel $\chi$, we may decompose the process into the three steps, and the above results demonstrate that $F$ is nondecreasing in each of the three steps. Therefore, Eq. (1.33) holds.

**No-cloning theorem.** An immediate consequence of the nondecreasing property of the fidelity is the no-cloning theorem. Consider a cloning machine that would transform an arbitrary input pure state $\hat{\rho}_{\mathrm{in},\phi} := |\phi\rangle_{AA}\langle\phi|$ into a duplicated pure state $\hat{\rho}_{\mathrm{out},\phi} := |\phi\rangle_{AA}\langle\phi| \otimes |\phi\rangle_{A'A'}\langle\phi|$. For $0 < |{}_A\langle\phi|\psi\rangle_A|^2 < 1$, we would have

$$F(\hat{\rho}_{\mathrm{out},\phi}, \hat{\rho}_{\mathrm{out},\psi}) = F(\hat{\rho}_{\mathrm{in},\phi}, \hat{\rho}_{\mathrm{in},\psi})^2 < F(\hat{\rho}_{\mathrm{in},\phi}, \hat{\rho}_{\mathrm{in},\psi}), \tag{1.34}$$

which violates Eq. (1.33). Therefore, this cloning machine could never exist.

## 1.5 Communication Resources

The task of sending quantum information is essentially different from that of sending classical information, and is achieved using a dedicated quantum channel. Interestingly, transmission of quantum information can also be achieved by supplementing a classical channel with another resource: entanglement. In this subsection, we will see how the three communication resources are related to each other, while focusing our discussion on the ideal cases.

### 1.5.1 Quantum Channels and Classical Channels

An *ideal classical channel* will transmit a symbol chosen from a fixed set $\{1, 2, \ldots, d\}$ without any error from a sender to a receiver. The number of symbols $d$ stands for the usefulness of the channel as a resource. A channel with $d = 2$ is normally regarded to have a unit of usefulness, called a *bit*. General ideal channels with $d$ symbols have $\log_2 d$ bits. This makes sense because the combined use of a ($\log_2 d$)-bit channel and a ($\log_2 d'$)-bit channel amounts to the single use of a ($\log_2 d + \log_2 d'$)-bit channel.

In a similar vein, we consider an *ideal quantum channel*, which faithfully transmits the arbitrary quantum states of a $d$-level physical system that is associated with a Hilbert space of dimension $d$. Because we have already called the two-level system a qubit, let us define the usefulness of such a channel as ($\log_2 d$) *qubits*. Because $\dim(\mathscr{H} \otimes \mathscr{H}') = (\dim \mathscr{H})(\dim \mathscr{H}')$, this measure is additive for the combined use of ideal channels.

We now consider how the two types of channels differ. First, a quantum channel can never be simulated using any amount of classical channels. This is because of the no-cloning theorem, as described in Sect. 1.4.4. Because the output of a classical channel can be freely copied, if the receiver were able to reconstruct any input state $|\phi\rangle$, then they could repeat the same procedure to create another copy of state $|\phi\rangle$, which is forbidden by the no-cloning theorem.

In contrast, a ($\log_2 d$)-qubit quantum channel can be used to simulate a classical channel. To simulate a ($\log_2 d'$)-bit channel, the sender can encode a symbol $i \in \{1, 2, \ldots, d'\}$ on a quantum state, i.e., the sender transmits the quantum state $\hat{\rho}_i$ via the quantum channel, according to the symbol $i$ that is to be transmitted. The receiver can then perform a measurement of the transmitted state to decode the index $i$. Encoding on mutually orthogonal states certainly works if $d' = d$, but the user may want to exploit the fact that there are an infinite number of different quantum states to transmit larger numbers of symbols. To deny any such possibility, we recall that any measurement strategy must be described as in Eq. (1.29), using a POVM $\{\hat{F}_j\}$. To simulate an ideal channel, $\mathrm{Tr}(\hat{F}_i \hat{\rho}_i) = 1$ should hold for $i = 1, \ldots, d'$. Because $\{\hat{F}_j\}$ are positive and $\sum_{i=1}^{d'} \hat{F}_i \leq \hat{1}$, we have $d' = \sum_{i=1}^{d'} \mathrm{Tr}(\hat{F}_i \hat{\rho}_i) \leq \sum_{i=1}^{d'} \mathrm{Tr}(\hat{F}_i) \leq \mathrm{Tr}\hat{1} = d$, thus proving the following.

**Theorem 4.** *Without use of another communication resource, a* ($\log_2 d$)*-qubit ideal quantum channel can never simulate a* ($\log_2 d'$)*-bit ideal classical channel if* $d' > d$.

### 1.5.2 Entanglement as a Communication Resource

We have seen that a quantum channel is qualitatively different from a classical channel. We may then ask what exactly is the difference between the channels, or ask what kind of communication resources may be used to complement a classical channel to enable it to simulate a quantum channel. It turns out that the entanglement is the answer to these questions.

As an ideal resource of entanglement, let us consider a maximally entangled state with a Schmidt number of $d$,

$$|\Phi_{0,0}\rangle_{AB} := \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |j\rangle_A |j\rangle_B \tag{1.35}$$

where $\{|j\rangle_A\}$ and $\{|j\rangle_B\}$ are the orthonormal bases of $\mathscr{H}_A$ and $\mathscr{H}_B$, respectively. When each subsystem is held by the sender and by the receiver, we can quantify the usefulness of this state as $(\log_2 d)$ *ebits*, which is additive when two or more maximally entangled states are available. Any state that is written as $(\hat{U}_A \otimes \hat{V}_B)|\Phi_{0,0}\rangle_{AB}$ is also a maximally entangled state and is regarded as a resource of the same number of ebits.

If a $(\log_2 d)$-qubit quantum channel is available, then the sender can create state $|\Phi_{0,0}\rangle_{AB}$ locally and transmit system $B$ to the receiver, which produces $(\log_2 d)$ ebits of entanglement resource.

**Theorem 5 (Entanglement sharing).** *A* $(\log_2 d)$*-qubit ideal quantum channel can be converted into* $(\log_2 d)$ *ebits of ideal entanglement.*

Next, let us compare entanglement with classical channels. First, entanglement does not help in augmentation of a classical channel.

**Theorem 6.** *Without use of another communication resource, no amount of entanglement can convert a* $(\log_2 d)$*-bit ideal classical channel into a* $(\log_2 d')$*-bit ideal classical channel with* $d' > d$*.*

*Proof.* Suppose that the sender chooses a symbol $i \in \{1, 2, \ldots, d'\}$ at random. Assume that it is possible to transmit $i$ faithfully by using a $(\log_2 d)$-bit ideal classical channel and shared entanglement. Because the output of the channel can be guessed correctly with a probability of $1/d$ by random guessing, the receiver can form a strategy, which, without communication, allows the symbol $i$ to be guessed with a success probability of $1/d$. Therefore, $1/d \le 1/d'$ must hold. □

Entanglement is a *static* resource in the sense that it is simply a correlation and it does not refer to any transfer of information. A classical channel is *dynamic* with regard to its ability to move information around. In this respect, the theorem above may be regarded as a natural example where a static resource cannot be converted into a dynamic resource. However, there is a subtlety here that will be manifest when we see the protocol for quantum dense coding in Sect. 1.5.4.

Finally, we consider the reverse question of how entanglement can be manipulated with unlimited use of classical channels. Suppose that Alice and Bob can freely use classical channels between them in both directions, and they can locally perform any physically allowed measurement or state transformation. This type of framework is called *LOCC (local operations and classical communication).*

Suppose that Alice and Bob initially share a pure bipartite state $|\Psi\rangle_{AB}$, and try to transform this state into other states under the LOCC framework. Without loss of generality, we may assume that only one party is conducting a local operation at any one time. This means that Alice first conducts a local operation, reveals an outcome to Bob through a classical communication, and Bob then conducts a local operation in turn, and so on. For Alice's turn, her operation is generally written as in Eq. (1.28), with $\hat{M}^{(j,k)}$ acting on Alice's system alone. Although the general description includes the index $k$, which is discarded, for the purposes of state transformation, Alice may as well record this index. Therefore, we omit $k$ and conclude that, after Alice's first turn, Alice and Bob share state $|\Psi^{(j)}\rangle_{A'B}$ with probability $p_j$, where

$$p_j|\Psi^{(j)}\rangle_{A'BA'B}\langle\Psi^{(j)}| = (\hat{M}_A^{(j)} \otimes \hat{1}_B)|\Psi\rangle_{ABAB}\langle\Psi|(\hat{M}_A^{(j)} \otimes \hat{1}_B)^\dagger \tag{1.36}$$

and $\hat{M}_A^{(j)} : \mathscr{H}_A \to \mathscr{H}_{A'}$ satisfies $\sum_j \hat{M}_A^{(j)\dagger}\hat{M}_A^{(j)} = \hat{1}_A$. Let $\hat{\rho}_B$ and $\hat{\rho}_B^{(j)}$ be the marginal density operators of system B for $|\Psi\rangle_{AB}$ and $|\Psi^{(j)}\rangle_{A'B}$, respectively. Taking a partial trace and summation over $j$ in Eq. (1.36), we have

$$\sum_j p_j \hat{\rho}_B^{(j)} = \hat{\rho}_B. \tag{1.37}$$

This equation shows that the rank of $\hat{\rho}_B^{(j)}$ never exceeds that of $\hat{\rho}_B$. The Schmidt number of state $|\Psi^{(j)}\rangle_{A'B}$ therefore never exceeds that of the initial state $|\Psi\rangle_{AB}$. A similar argument is applicable to Bob's turns, and we thus see that the Schmidt number never increases under the LOCC framework, even probabilistically. Specifically, no entanglement is generated under the LOCC framework when starting from a product state with a Schmidt number of unity. This is often adopted as a defining property of entanglement when discussing more general cases of mixed-state entanglement.

In view of the relationships between the communication resources, the above argument means that the classical channels do not help to increase entanglement, and this is summarized as follows.

**Theorem 7.** *Without use of another communication resource, no amount of communication over classical channels can convert a* $(\log_2 d)$*-ebit ideal entanglement into a* $(\log_2 d')$*-ebit ideal entanglement with* $d' > d$.

This theorem implies that entanglement has a nonclassical aspect that cannot be replaced by classical channels. If we combine the two resources, we will obtain a resource that is both dynamic and nonclassical, and we may perhaps simulate a quantum channel. This is indeed true, and will be explained in Sect. 1.5.4 after we summarize the properties of the maximally entangled states in Sect. 1.5.3.

### 1.5.3 Properties of Maximally Entangled States

Let $\mathscr{H}_A$ and $\mathscr{H}_B$ be Hilbert spaces of dimension $d$ for the systems $A$ and $B$. Here, we summarize the relevant properties of the maximally entangled states of system $AB$.

(E1) All maximally entangled states have a common marginal state $d^{-1}\hat{1}_A$ for subsystem $A$, and a common marginal state $d^{-1}\hat{1}_B$ for subsystem $B$.

(E2) For any pair of maximally entangled states $|\Phi\rangle_{AB}$ and $|\Phi'\rangle_{AB}$, unitary operators $\hat{U}_A$ and $\hat{V}_B$ exist such that $|\Phi'\rangle_{AB} = (\hat{U}_A \otimes \hat{1}_B)|\Phi\rangle_{AB} = (\hat{1}_A \otimes \hat{V}_B)|\Phi\rangle_{AB}$.

(E3) A maximally entangled state $|\Phi\rangle_{AB}$ specifies a one-to-one correspondence $|\phi\rangle_A \leftrightarrow |\phi^*\rangle_B$ between the pure states of subsystem $A$ and those of subsystem $B$, as characterized by $d^{-1/2}|\phi\rangle_A = {}_B\langle\phi^*||\Phi\rangle_{AB}$ and $d^{-1/2}|\phi^*\rangle_B = {}_A\langle\phi||\Phi\rangle_{AB}$.

(E4) A maximally entangled state $|\Phi\rangle_{AB}$ specifies a one-to-one correspondence $\hat{M}_A \leftrightarrow \hat{M}_B^{\mathrm{T}}$ between the operators that act on $\mathscr{H}_A$ and those acting on $\mathscr{H}_B$, as characterized by

$$(\hat{M}_A \otimes \hat{1}_B)|\Phi\rangle_{AB} = (\hat{1}_A \otimes \hat{M}_B^{\mathrm{T}})|\Phi\rangle_{AB}. \tag{1.38}$$

Specifically, if $\hat{M}_A$ is unitary then $\hat{M}_B^{\mathrm{T}}$ is also unitary, and vice versa.

(E5) There is an orthonormal basis $\{|\Phi_{l,m}\rangle_{AB}\}_{l=0,\ldots,d-1}^{m=0,\ldots,d-1}$ of $\mathscr{H}_A \otimes \mathscr{H}_B$ where every basis state is a maximally entangled state. This type of basis is called a *Bell* basis.

(E1) is the definition given in Sect. 1.2.5. (E2) is a combination of (E1) and Theorem 3. (E3) refers to the relative states explained in Sect. 1.4.2.

For (E5), a Bell basis that includes state $|\Phi_{0,0}\rangle_{AB}$ of Eq. (1.35) is constructed as follows. For each subsystem, we define unitary operators

$$\hat{X} := \sum_{j=0}^{d-1} |j+1 \pmod{d}\rangle\langle j| \quad \text{and} \quad \hat{Z} := \sum_{j=0}^{d-1} \beta^j |j\rangle\langle j| \tag{1.39}$$

with $\beta := \exp(2\pi i/d)$. Using these operators, we define $|\Phi_{l,m}\rangle_{AB} := (\hat{X}_A^l \otimes \hat{Z}_B^m)|\Phi_{0,0}\rangle_{AB}$. Using the relation $\hat{Z}\hat{X} = \beta\hat{X}\hat{Z}$, it is simple to show that $|\Phi_{l,m}\rangle_{AB}$ is a simultaneous eigenvector of the commuting unitary operators $\hat{X}_A \otimes \hat{X}_B$ and $\hat{Z}_A \otimes \hat{Z}_B^{-1}$ with eigenvalues of $\beta^{-m}$ and $\beta^l$, respectively. Therefore, the $d^2$ states $\{|\Phi_{l,m}\rangle_{AB}\}_{l=0,\ldots,d-1}^{m=0,\ldots,d-1}$ are all orthogonal. For $d = 2$, the Bell basis consists of the following states.

$$|\Phi_+\rangle = |\Phi_{0,0}\rangle = 2^{-1/2}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B) \tag{1.40}$$

$$|\Phi_-\rangle = |\Phi_{0,1}\rangle = 2^{-1/2}(|0\rangle_A|0\rangle_B - |1\rangle_A|1\rangle_B) \tag{1.41}$$

$$|\Psi_+\rangle = |\Phi_{1,0}\rangle = 2^{-1/2}(|1\rangle_A|0\rangle_B + |0\rangle_A|1\rangle_B) \tag{1.42}$$

$$|\Psi_-\rangle = |\Phi_{1,1}\rangle = 2^{-1/2}(|1\rangle_A|0\rangle_B - |0\rangle_A|1\rangle_B) \tag{1.43}$$

(E4) is confirmed as follows. Suppose that $|\Phi\rangle_{AB}$ is decomposed as shown in Eq. (1.7). By applying ${}_A\langle u_i|_B\langle v_j|$ to Eq. (1.38), we see that Eq. (1.38) is equivalent to

$$ {}_A\langle u_i|\hat{M}_A|u_j\rangle_A = {}_B\langle v_j|\hat{M}_B^{\mathrm{T}}|v_i\rangle_B \tag{1.44}$$

for $i, j = 1, \ldots, d$. This means that the matrix representation of $\hat{M}_B^{\mathrm{T}}$ in the basis $\{|v_i\rangle_B\}$ is the transpose of the matrix representation of $\hat{M}_A$ in the basis $\{|u_i\rangle_A\}$.

As an example of Property (E4), the following relations are worth mentioning:

$$(\hat{X}_A \otimes \hat{1}_B)|\Phi_{0,0}\rangle_{AB} = (\hat{1}_A \otimes \hat{X}_B^{-1})|\Phi_{0,0}\rangle_{AB} \tag{1.45}$$

$$(\hat{Z}_A \otimes \hat{1}_B)|\Phi_{0,0}\rangle_{AB} = (\hat{1}_A \otimes \hat{Z}_B)|\Phi_{0,0}\rangle_{AB}, \tag{1.46}$$

and can easily be confirmed.

### *1.5.4 Quantum Dense Coding and Quantum Teleportation*

In this subsection, we explain two types of scheme in which shared entanglement helps with the conversion between the quantum and classical channels. Every subsystem $X$ that appears in this subsection is a $d$-level system with Hilbert space of dimension $d$, and with a standard orthonormal basis denoted by $\{|j\rangle_X\}$. The Bell basis is defined for each pair of subsystems according to the standard bases.

In Theorem 4, we have seen that a one-qubit quantum channel alone can only send one bit of classical information. If the sender and the receiver share entanglement beforehand, then the quantum channel can send more via a protocol called *quantum dense coding* [8].

**Theorem 8 (Quantum dense coding).** *A* $(\log_2 d)$*-qubit ideal quantum channel and a* $(\log_2 d)$*-ebit ideal entanglement can be converted into a* $(2\log_2 d)$*-bit ideal classical channel.*

A protocol for quantum dense coding can be constructed simply by using the Bell basis $\{|\Phi_{l,m}\rangle_{AB}\}$ of the two $d$-level subsystems, i.e., Property (E5) in Sect. 1.5.3. We show that Alice can send Bob a symbol $(l, m)$ that was chosen from $d^2$ candidates $\{(l,m)\}_{l=0,\ldots,d-1}^{m=0,\ldots,d-1}$. Suppose that Alice and Bob shared the entangled state $|\Phi_{0,0}\rangle_{AB}$ initially. Property (E2) ensures that Alice can locally transform[6] the state $|\Phi_{0,0}\rangle$ into the state $|\Phi_{l,m}\rangle$ that is specified by the chosen symbol $(l, m)$. She then sends

[6] An explicit form of Alice's transformation is $|\Phi_{l,m}\rangle_{AB} = (\hat{X}_A^l \hat{Z}_A^m \otimes \hat{1}_B)|\Phi_{0,0}\rangle_{AB}$, which is obtained from Eq. (1.46).

subsystem $A$, which has a Hilbert space with dimension $d$, to Bob using the $(\log_2 d)$-qubit quantum channel. Bob, who now holds both subsystems $A$ and $B$, conducts an orthogonal measurement with the Bell basis $\{|\Phi_{l,m}\rangle_{AB}\}$ to determine Alice's choice $(l, m)$.

This protocol is remarkable in the sense that the static resource of entanglement enhances an ideal channel's ability to achieve the dynamic task of information transmission. This is in stark contrast with what we saw in Theorem 6, i.e., that the static resource of entanglement cannot augment the dynamic resources of classical channels.

Next, we explain the protocol of *quantum teleportation* [9], which combines the nonclassical resource of entanglement and the dynamic resource of a classical channel to achieve faithful transmission of quantum states.

**Theorem 9 (Quantum teleportation).** *A* $(2\log_2 d)$*-bit ideal classical channel and a* $(\log_2 d)$*-ebit ideal entanglement can be converted into a* $(\log_2 d)$*-qubit ideal quantum channel.*

The protocol proceeds as follows. Suppose that Alice and Bob initially share the entangled states $|\Phi_{0,0}\rangle_{AB}$ of two $d$-level systems. Alice also holds another $d$-level subsystem $A'$, and she is supposed to transmit the state of this subsystem to Bob. Alice first performs an orthogonal measurement with the Bell basis $\{|\Phi_{l,m}\rangle_{AA'}\}$ on subsystems $A$ and $A'$, and transmits the outcome $(l, m)$ to Bob through the $(2\log_2 d)$-bit classical channel. Based on the received indices $(l, m)$, Bob then applies a unitary transformation $\hat{U}_B^{(l,m)}$ to subsystem $B$.

We now consider how we can choose $\hat{U}_B^{(l,m)}$ such that the final state of system $B$ is always identical to the initial state of system $A'$ (see also Fig. 1.5). Consider another $d$-level system $R$, and suppose that the system $A'R$ is initially prepared

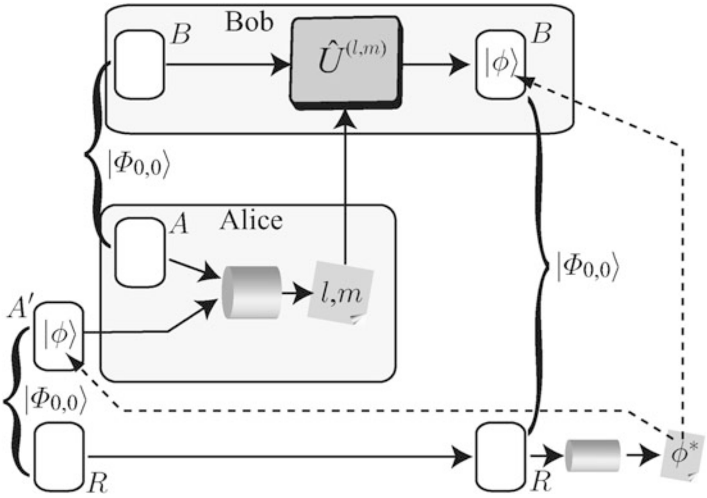

**Fig. 1.5** Entanglement swapping and quantum teleportation

in state $|\Phi_{0,0}\rangle_{A'R}$. Later, at the end of this argument, we will use Property (E3) to discuss the case where $A'$ is initially prepared in the general state $|\phi\rangle_{A'}$.

We begin with the following relation, which can be easily confirmed from the definition of Eq. (1.35):

$$d^{-1}|\Phi_{0,0}\rangle_{BR} = {}_{AA'}\langle\Phi_{0,0}||\Phi_{0,0}\rangle_{AB}|\Phi_{0,0}\rangle_{A'R}. \tag{1.47}$$

According to Theorem 1, this shows that if the outcome is $(l, m) = (0, 0)$, then the state of the system $BR$ is $|\Phi_{0,0}\rangle_{BR}$. We want to generalize this relationship to the case where ${}_{AA'}\langle\Phi_{0,0}|$ is replaced by ${}_{AA'}\langle\Phi_{l,m}|$. From Eq. (1.46), we have $|\Phi_{l,m}\rangle_{AA'} = (\hat{X}_A^l\hat{Z}_A^m \otimes \hat{1}_{A'})|\Phi_{0,0}\rangle_{AA'}$ and thus ${}_{AA'}\langle\Phi_{l,m}| = {}_{AA'}\langle\Phi_{0,0}|(\hat{Z}_A^{-m}\hat{X}_A^{-l} \otimes \hat{1}_{A'})$. From Eqs. (1.45) and (1.46), we have $(\hat{Z}_A^{-m}\hat{X}_A^{-l} \otimes \hat{1}_B)|\Phi_{0,0}\rangle_{AB} = (\hat{1}_A \otimes \hat{X}_B^l\hat{Z}_B^{-m})|\Phi_{0,0}\rangle_{AB}$. We therefore obtain

$$d^{-1}(\hat{X}_B^l\hat{Z}_B^{-m} \otimes \hat{1}_R)|\Phi_{0,0}\rangle_{BR} = {}_{AA'}\langle\Phi_{l,m}||\Phi_{0,0}\rangle_{AB}|\Phi_{0,0}\rangle_{A'R}, \tag{1.48}$$

which identifies the state of the system $BR$ after the Bell measurement in the protocol. By setting $\hat{U}_B^{(l,m)} = \hat{Z}_B^m\hat{X}_B^{-l}$, the protocol should leave the system $BR$ in the same state, $|\Phi_{0,0}\rangle_{BR}$, regardless of the value of the outcome $(l, m)$. In summary, if we begin with state $|\Phi_{0,0}\rangle_{A'R}$, the protocol then transforms it into state $|\Phi_{0,0}\rangle_{BR}$, in which the system with which $R$ is entangled changes from $A'$, possessed by Alice, to $B$, which is held by Bob. This procedure is often called *entanglement swapping* [10].

The case where system $A'$ is initially prepared in the arbitrary state $|\phi\rangle_{A'} = \sum_j c_j|j\rangle_{A'}$ can be analyzed using Property (E3) on the relative states, as in Sect. 1.4.2. After entanglement swapping, the state of the system $BR$ is $|\Phi_{0,0}\rangle_{BR}$. Suppose that we perform an orthogonal measurement on system $R$ with a basis that includes a state $|\phi^*\rangle_R = \sum_j \bar{c}_j|j\rangle_R$. If the corresponding outcome is obtained, then the state of system $B$ becomes its relative state, $|\phi\rangle_B = \sum_j c_j|j\rangle_B$. Because the entanglement swapping protocol starts with state $|\Phi_{0,0}\rangle_{A'R}$ and does not operate on system $R$, Theorem 1 then dictates that such an event must be consistent with the case where system $A'$ was initially prepared in $|\phi\rangle_{A'} = \sum_j c_j|j\rangle_{A'}$. We thus conclude that if we carry out the protocol with initial state $|\phi\rangle_{A'} = \sum_j c_j|j\rangle_{A'}$, the final state of system $B$ is $|\phi\rangle_B = \sum_j c_j|j\rangle_B$, which is regarded as a faithful transmission of the quantum state.

The existence of this quantum teleportation protocol has profound consequences. Because classical channels are much easier to implement in practice, let us assume that these channels can be used freely in both directions between Alice and Bob. The quantum teleportation protocol and the entanglement sharing of Theorem 5 then imply that one qubit of dynamic resource and one ebit of static resource are freely interconvertible. Because the static resource of entanglement can be stored in quantum memories, this effectively allows dynamic resource storage. If the quantum channels are not ideal but noisy, we may convert these channels into noisy entanglement, which is then distilled into close-to-ideal entanglement and can be used for faithful quantum transmission. When we wish to concatenate the quantum channels, which will only work probabilistically, as in the case of transmission of photons over an optical fiber, the combination of entanglement sharing and entanglement swapping dramatically improves the process efficiency, as described in Chap. 4. It should also be noted that entanglement has no preferred direction. We can convert a quantum channel from Alice to Bob into a channel from Bob to Alice, through the protocol of entanglement sharing followed by quantum teleportation with backward classical communication.

### *1.5.5 Conversion Among the Resources*

In the preceding subsections, we have described three protocols, entanglement sharing, quantum dense coding, and quantum teleportation, that provide conversion among the three types of communication resources: ebits, bits, and qubits. Because these protocols were introduced in a rather heuristic way, we might expect that there are many other protocols that can be used for resource conversion. Here, we argue that this is not the case. The three protocols in a sense exhaust all possibilities as far as conversion among the three ideal resource types is concerned.

Imagine that Alice and Bob have a right to use $E$ ebits of a shared ideal entanglement, $C$ bits of an ideal classical channel, and $Q$ qubits of an ideal quantum channel, which we denote by the portfolio $(E, C, Q)$. According to Theorems 5, 8, and 9, the three protocols change the portfolio in the following way.

| | |
|---|---|
| Entanglement sharing (ES) | $(E, C, Q) \rightarrow (E + 1, C, Q - 1)$ |
| Quantum dense coding (DC) | $(E, C, Q) \rightarrow (E - 1, C + 2, Q - 1)$ |
| Quantum teleportation (QT) | $(E, C, Q) \rightarrow (E - 1, C - 2, Q + 1)$ |

Let us assume that we start from $(E_0, C_0, Q_0)$. By repeating these protocols $N_{\mathrm{ES}}$, $N_{\mathrm{DC}}$, and $N_{\mathrm{QT}}$ times,

$$(E_0, C_0, Q_0) + N_{\mathrm{ES}}(1, 0, -1) + N_{\mathrm{DC}}(-1, 2, -1) + N_{\mathrm{QT}}(-1, -2, 1) \quad (1.49)$$

is attainable. Therefore, if we ignore the fact that only a discrete set of points is attainable, we may say that it is possible to reach anywhere within a triangular pyramid with apex $(E_0, C_0, Q_0)$ and with edges defined by the vectors $(1, 0, -1)$, $(-1, 2, -1)$, and $(-1, -2, 1)$ (see Fig. 1.6).

We are now interested in whether we can reach a point outside this pyramid. We have already derived various restrictions on resource conversion in Theorems 4, 6, and 7, which are summarized as follows.

| | |
|---|---|
| Theorem 4 | $(0, 0, Q) \rightarrow (0, C', 0)$ only if $C' \leq Q$. |
| Theorem 6 | $(E, C, 0) \rightarrow (E', C', 0)$ only if $C' \leq C$ |
| Theorem 7 | $(E, C, 0) \rightarrow (E', C', 0)$ only if $E' \leq E$ |

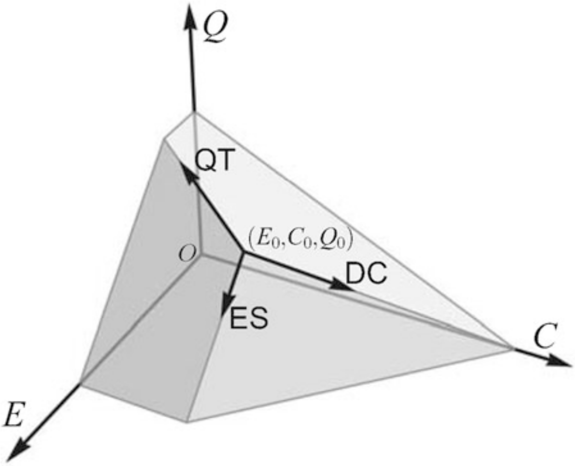

**Fig. 1.6** Permitted resource conversion region. *ES*: entanglement sharing; *DC*: quantum dense coding; *QT*: quantum teleportation

Using these theorems, we derive a restriction on a general protocol $\mathscr{P}$ that performs the conversion from $(E_0, C_0, Q_0) \rightarrow (E, C, Q)$. This is done by combining $\mathscr{P}$ with the three protocols, such that the theorems above are applicable to the entire conversion process. For example, we have

$$(0, 0, Q_0 + Q + 2C_0 + E_0 + E) \xrightarrow{\text{ES}} (Q + C_0 + E_0, 0, Q_0 + C_0 + E) \xrightarrow{\text{DC}}$$
$$(Q + E_0, 2C_0, Q_0 + E) \xrightarrow{\mathscr{P}} (Q + E, C_0 + C, Q + E) \xrightarrow{\text{DC}}$$
$$(0, 2Q + C_0 + C + 2E, 0),$$

which, from Theorem 4, requires that

$$(E - E_0) + (C - C_0) + (Q - Q_0) \leq 0. \tag{1.50}$$

Similarly, from

$$(Q_0 + Q + E_0, C_0 + 2Q_0, 0) \xrightarrow{\text{QT}} (Q + E_0, C_0, Q_0) \xrightarrow{\mathscr{P}} (Q + E, C, Q)$$
$$\xrightarrow{\text{DC}} (E, C + 2Q, 0),$$

we use Theorem 6 to obtain

$$(C - C_0) + 2(Q - Q_0) \leq 0. \tag{1.51}$$

Finally, by applying Theorem 7 to

$$(Q_0 + E_0, C_0 + 2Q_0, 0) \xrightarrow{\text{QT}} (E_0, C_0, Q_0) \xrightarrow{\mathscr{P}} (E, C, Q) \xrightarrow{\text{ES}} (Q + E, C, 0),$$

we have

$$(E - E_0) + (Q - Q_0) \leq 0. \tag{1.52}$$

It is simple to confirm that Eqs. (1.50), (1.51), and (1.52) correspond to the three faces of the pyramid in Fig. 1.6. It is thus impossible to reach any point outside the pyramid. We see that the three protocols of entanglement sharing, quantum dense coding, and quantum teleportation correspond to the three edges of the achievable region, and form a unique triad that governs the conversions that are allowed among the resources of quantum channels, classical channels, and entanglement.

## References

1. N. Gisin, Helv. Phys. Acta **62**(4), 363 (1989)
2. L.P. Hughston, R. Jozsa, W.K. Wootters, Phys. Lett. A **183**(1), 14 (1993)
3. H.K. Lo, H.F. Chau, Phys. Rev. Lett. **78**(17), 3410 (1997)
4. D. Mayers, Phys. Rev. Lett. **78**(17), 3414 (1997)
5. A. Peres, Phys. Lett. A **128**(1), 19 (1988)
6. A. Uhlmann, Rep. Math. Phys. **9**(2), 273 (1976)
7. R. Jozsa, J. Modern Opt. **41**(12), 2315 (1994)
8. C.H. Bennett, S.J. Wiesner, Phys. Rev. Lett. **69**(20), 2881 (1992)
9. C.H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, W.K. Wootters, Phys. Rev. Lett. **70**(13), 1895 (1993)
10. M. Żukowski, A. Zeilinger, M. Horne, A. Ekert, Phys. Rev. Lett. **71**(26), 4287 (1993)